

# Local identification of scalar hybrid models with tree structure

Bernold Fiedler\*, Andreas Schuppert\*\*

\* Freie Universität Berlin  
Institut für Mathematik I  
Arnimallee 2–6  
14195 Berlin, Germany

\*\* Bayer Technology Services GmbH  
PT - M&T - CS  
51368 Leverkusen, Germany

October 2004

## Abstract

Standard modeling approaches, for example in chemical engineering, suffer from two principal difficulties: the curse of dimension and a lack of extrapolability. We propose an approach via structured hybrid models to resolve both issues.

For simplicity we consider reactor models which can be written as a tree-like composition of scalar input-output functions  $u_j$ . The vertices  $j$  of the finite tree structure represent known or unknown sub-processes of the overall process. Known processes are modeled by white-box functions  $u_j$ ; unknown processes are represented by black boxes  $u_j$ . Oriented edges of the tree indicate composition of the input-output relations  $u_j$  in a feed forward structure. The tree structure of a mixture of black and white boxes constitutes what we call a *structured hybrid model* (SHM).

Under certain assumptions on differentiability, genericity, and monotonicity, we provide an inductive algorithm which uniquely identifies all black boxes in the SHM, up to a trivial scaling calibration between adjacent black boxes.

Our result does not require any extra measurements interior to the SHM. Instead, we only require global, overall input-output data, clustered along a  $d$ -dimensional data base of inputs. The dimension  $d$  need not exceed the maximal input dimension of any individual black box in the SHM. Compared to the total input dimension of the reactor, which may be much higher than  $d$ , this dimension reduction effectively avoids the curse of dimension. Moreover, our unique identification of all black boxes accomodates a reliable global extrapolation, far beyond the original data base, to input-regions of full dimension.

We illustrate our results with a model of an industrial continuous polymerization plant.

## 1 Introduction

The mathematical framework of structured hybrid models originates from technological problems, for example in chemical engineering. The *curse of dimension* of data analysis is a principal motivation here. One of the most important bottlenecks

for the application of model based process optimization is in fact the typical lack of reliable, quantitative models  $u_j$  for reactive subprocesses. Therefore advanced modeling techniques allowing the access to new areas of applications promise substantial economic potential. The classical modeling techniques suffer from several shortcomings.

Rigorous models based on quantitative scientific principles require detailed knowledge of all subprocesses  $u_j$  involved in the process. Although highly desirable, it is often not affordable to close the gaps in quantitative process knowledge for all subprocesses, and thereby allow rigorous modeling.

Black-box modeling techniques on the other hand, like neural networks, multivariate splines, radial basis functions, or any other approximation method, allow the quantitative description of the *input-output* (i/o) *relations* of the overall process, which is often sufficient for industrial use. Such black-box techniques, however, often fail in process modeling applications because of two problems:

- (i) A high number of statistically well distributed data sets in the input variables is required for modeling of complex processes (*curse of dimension*). Typical production data, however, are highly correlated. Therefore the number of data sets available from comparable production conditions is usually small, clustered along low-dimensional subsets of input space.
- (ii) Due to the lack of extrapolability, black-box models cannot be applied for optimization.

These shortcomings in rigorous as well as in black-box modeling have led to the development of modeling techniques which combine both approaches [20, 16, 6, 8]. A significant improvement over black-box methods is achieved by the additional use of a-priori known structural information about the process. The central idea is the use of structural information to reduce the complexity of the black-box submodels  $u_j$ . Due to a reduction in complexity, compared with pure black-box models, the number of data sets required to identify the model can be reduced significantly without any loss of accuracy. This approach is called “*Structured Hybrid Modeling*”. A structured hybrid model (SHM) therefore consists of three components:

- (a) rigorous submodels  $u_j$ ,  $j \in W$ , describing the i/o-relation of those subprocesses which are well understood (white boxes)
- (b) black-box submodels for those subprocesses  $u_j$ ,  $j \in B$ , for which no rigorous models are available
- (c) a model flowsheet  $\Gamma$  describing the i/o-structure of all submodels, according to the real process.

The new ingredient in the SHM approach is the explicit use of the process structure via the model flowsheet. This incorporates additional qualitative knowledge and significantly reduces the complexity of the model. The main improvement, however, is not a mere reduction of complexity. The explicit use of the process flowsheet structure  $\Gamma$  in fact results in a drastical improvement of the performance of the model as well as in the ability to extrapolate far beyond ill-distributed training data sets. These features have first been described in the framework of reaction-kinetics modeling [17] and will be illustrated in the present paper with an example of a continuous polymerization process; see section 6.

Mathematically speaking, by a *structured hybrid model* (SHM) we mean a finite oriented graph  $\Gamma = \{1, \dots, N\}$  equipped with certain functions  $u_j$ , associated to the vertices  $j \in \Gamma$ , and with *global input-output data*  $h$ . Edges pointing towards vertex  $j$  are called *inputs* of  $u_j$ , whereas edges pointing away from  $j$  are called *outputs*. The oriented edges of the graph  $\Gamma$  indicate composition of the vertex functions  $u_j$ , thus giving rise to  $h$ . In other words, let an oriented edge point from vertex  $i$  to vertex  $j$ . We then substitute the output function  $u_i$  for the input argument  $x_i$  of the function  $u_j$ . Completing this substitution process over the graph  $\Gamma$ , we obtain the global input-output data  $h$  as the resulting composition of functions  $u_j$ ,  $j \in \Gamma$ . The functions  $u_j$  are considered as given for  $j$  in a set  $W \subset \Gamma$  of *white boxes*. The identification task requires to determine the functions  $u_j$  in a set  $j \in B \subset \Gamma$  of *black boxes*, disjoint from the white-box set  $W$ . The identification should be based on minimal information on the global input-output data  $h$ .

We call the above setting a structured hybrid model, first, because it possesses the *structure* of an oriented graph. Second, the global model of the data  $h$  is given by the composition of the vertex functions  $u_j$ , which themselves may *model* certain

biological, chemical, physical, economic, social, or other processes. Finally, our global model is of *hybrid* type, mixing both well-modeled processes  $u_j$  of the white boxes  $j \in W$  and completely unknown processes  $u_j$  of the black boxes  $j \in B$ .

In the present paper, we show how minimal information on the global input-output data  $h$  can be sufficient to uniquely identify all black boxes. Moreover, we indicate how all black boxes can be identified, successively and constructively.

We only consider the very special case when the structure graph  $\Gamma$  is a tree, in the present paper. For a mathematically complete description of this setting see ((1.3) – (1.6)) below. We also restrict attention to scalar-valued black and white boxes, with strictly monotone dependence on their input components. In addition, we impose further generic nondegeneracy conditions.

Results on the vector-valued case, branched graphs  $\Gamma$ , and even certain recycle graphs are work in progress, at this stage. See the technical report [11] for a summary. We briefly comment on some chemical aspects of our monotonicity assumption here. Mathematically restrictive as this condition may seem, it is rather the identification of quantitative reaction rates for complex reactions, mostly unknown black boxes in their detailed mechanisms, which constitute the bottleneck in chemical engineering applications. The identification of those black boxes therefore closes the most important bottleneck for this important class of applications. Admittedly many reaction mechanisms are highly nonlinear with quantitatively unknown nonlinearities. The trend in the reaction rates, when the concentration of an educt is increased, may be decreasing if the educt acts as an inhibitor, or else increasing for an activator. The sign of the effected change, however, alias the activator or inhibitor role played by the educt, remains constant in many cases. Therefore many reaction rates show monotone behavior on a detailed reaction level. This monotonicity feature is ubiquitous, not only in chemical models. It equally appears in metabolism modeling of biological processes, for example based on compositions of monotone Michaelis-Menten reactions. Our present monotonicity restriction is therefore much less restrictive from an application view point than it may appear, mathematically.

Early mathematical fame has been associated with a variant of our hybrid modeling problem. In problem XIII of his celebrated list, Hilbert has asked to express the roots of higher order polynomials by compositions of scalar functions with at most

two scalar input variables; [13]. In other words, the graph  $\Gamma$  was left unspecified (except for being a tree, binary at most), and all boxes were black, to be determined depending on the degree of the polynomial. In the class  $C^0$  of merely continuous functions, this problem was solved by Arnold and Kolmogorov, [2], [15]. In fact Kolmogorov observed that any function  $h \in C^0(\mathbb{R}^n, \mathbb{R})$  of  $n$  input variables can be written in the form

$$h(x_1, \dots, x_n) = \sum_{i=1}^{2n+1} u_i \left( \sum_{j=1}^n \psi_{i,j}(x_j) \right). \quad (1.1)$$

Note that the structure graph  $\Gamma$  associated to (1.1) is not a tree. The  $\psi_{i,j} \in C^0(\mathbb{R}, \mathbb{R})$  are universally prescribed given  $C^0$ -functions, hence white boxes. The scalar black-box functions  $u_j \in C^0(\mathbb{R}, \mathbb{R})$  can then be chosen such that (1.1) holds. Analogous results hold in the category of Lipschitz continuous functions [12].

In the class  $C^0$  these results illustrate a fundamental obstacle to our identification problem. The white-box functions  $\psi_{i,j}$  can be universally prescribed, and then fixed, according to a rather ingenious construction principle. There are many functions, however, which satisfy the conditions of the Arnold-Kolmogorov construction. Therefore unique identification is doomed to fail, in the  $C^0$  category, whenever we consider all  $u_j$  and at least one of the  $\psi_{i,j}$  in the Kolmogorov model (1.1) as a black box.

For related results in the category of functions  $h \in C^\nu(\mathbb{R}^n, \mathbb{R})$ ,  $u_j \in C^{\nu'}(\mathbb{R}^{n'}, \mathbb{R})$  which are  $\nu$  and  $\nu'$  times differentiable, respectively; see [22], [23]. These results rely on the *Vitushkin complexity*  $\nu/n$ . In fact, most data  $h$  cannot be represented by any finite composition of functions  $u_j$  of higher Vitushkin complexity

$$\nu'/n' > \nu/n. \quad (1.2)$$

Since  $\nu' = 0$  is not viable for unique identification purposes, by the Arnold-Kolmogorov result sketched above, we have to consider  $\nu' \geq 1$ . High Vitushkin complexity then arises for  $n' = 2$  inputs of the boxes  $u_j$ , for example. If the global input-output data  $h$  possess regularity  $\nu < n/2$ , the Vitushkin result therefore implies that most input-output data  $h$  cannot be realized by any choice of the structure graph  $\Gamma$  and of the boxes  $u_j$ . This can be interpreted as an instability result: arbitrarily small perturbations of  $h$  in the  $C^\nu$  topology will destroy identifiability of  $u_j$ , pushing any

solutions into non-existence. Conversely, *stable identification will necessarily require high differentiability of the data  $h$* . Although our uniqueness result in theorem 1.1 below does not address the stability issue, our completely explicit approach will reflect this unpleasant differentiability constraint.

We emphasize that the finite structure graph  $\Gamma$  is not prescribed a priori, in Vitushkin's results, but can be chosen to depend on the global input-output function  $h$  which is to be modeled. In hybrid modeling, in contrast, the graph  $\Gamma$  is prescribed. Moreover, the given global input-output function  $h$  is already known to be realized by the black- and white-box functions  $u_j$ . We recall our task: determine the black-box functions  $u_j$  from minimal information on  $h$ .

In the present paper, we restrict our attention to structure graphs  $\Gamma$  which are trees. In other words,  $\Gamma$  is connected via the edges, and does not possess (unoriented) cycles. For trees  $\Gamma$  we now repeat, in technically precise and complete terms, our definition of structured hybrid models, as given above. Specifically, because  $\Gamma$  is a tree, there exists a map

$$\begin{aligned} s : \Gamma \setminus \{N\} &\rightarrow \Gamma, \\ j &\mapsto s(j) \end{aligned} \tag{1.3}$$

which allocates to  $j$  its *successor*  $s(j)$ , pointing along the unique edge which leaves  $j$  in output direction. The successor function  $s$  decomposes the tree  $\Gamma$  into layers  $\ell_k$ ,  $k = 0, \dots, m$ , defined by

$$\begin{aligned} \ell_0 &:= \Gamma \setminus \text{range}(s), \\ \ell_k &:= s^k(\ell_0). \end{aligned} \tag{1.4}$$

The layer  $\ell_0$  will indicate the *inputs* to the structured hybrid model. Note that the final output  $\ell_m = \{N\}$  is the root of the tree  $\Gamma$ . We also use the notation  $\Gamma' := \Gamma \setminus \ell_0 = \text{range}(s)$  for the non-input vertices, decomposed disjointly into black and white boxes,

$$\Gamma' = B \cup W. \tag{1.5}$$

For  $j \in \Gamma'$ , we let  $p_j := s^{-1}(j)$  denote the *predecessors* of vertex  $j$ . The number  $|p_j|$  of inputs to vertex  $j$  may be any positive integer, depending on  $j$ .

To each black- or white-box vertex  $j \in \Gamma'$  we associate a scalar,  $\nu$  times differentiable function  $u_j \in C^\nu(\mathbb{R}^{|p_j|}, \mathbb{R})$  modeling the input-output relation at this box. We also

introduce variables  $x_j$ ,  $j \in \Gamma$ , to denote the *output* of box  $j$ , if  $j \in \Gamma'$ . If  $j \in \ell_0$  is at the entry layer then  $x_j$  denotes an input component to the hybrid model. We abbreviate by

$$x^0 := (x_j)_{j \in \ell_0} \quad \text{and} \quad x^j := (x_i)_{i \in p_j} \quad (1.6)$$

the *global input vector* and the *input vector* to box  $j$ , respectively. With this notation the composition of outputs and inputs along the tree structure  $\Gamma$  of the hybrid model takes the recursive form

$$x_j = u_j(x^j) = u_j((x_i)_{i \in p_j}), \quad (1.7)$$

for  $j \in \Gamma' = B \cup W$ . We recall that  $p_j = s^{-1}(j)$  denotes the predecessors of box  $j \in \Gamma'$ . For given global input vector  $x^0 = (x_i)_{i \in \ell_0}$  the relation (1.7) indeed defines all outputs, recursively for  $j$  in increasing layers  $\ell_k$ ,  $k = 1, \dots, m$  of the tree  $\Gamma$  until finally

$$x_N = h(x^0). \quad (1.8)$$

The final output  $x_N$  of the last box  $j = N$  can be seen, holistically, as a function of the input data  $x^0$  and a nonlinear functional  $\Phi^\Gamma$ , depending on the graph  $\Gamma$ , of the black- and white-box functions  $u_j$ :

$$\Phi^\Gamma((u_j)_{j \in B \cup W}, x^0) = x_N = h(x^0). \quad (1.9)$$

Our task is therefore to identify  $(u_j)_{j \in B}$  from (1.9), under minimal information on  $h$ . We call the above SHM *scalar* because  $h$  and all vertex functions  $u_j$  are scalar real-valued functions.

The ‘‘curse of dimension’’ constrains  $h$  to be evaluated along a set  $D$  of inputs  $x^0$  of the smallest possible dimension. We call  $D \subset \mathbb{R}^{|\ell_0|}$  a *data base* for the structured hybrid model (1.9), if  $D$  is a connected, embedded  $d$ -dimensional  $C^\nu$  submanifold of  $\mathbb{R}^{|\ell_0|}$ . Moreover we require the  $\nu$ -jet of the Taylor coefficients  $D^i h(x^0)$ ,  $|i| \leq \nu$ , to be given on  $x^0 \in D$ .

Consideration of the  $\nu$ -jet of the input-output data  $h$ , rather than just the values of  $h$  themselves, has been motivated by our above discussion of the Arnold-Kolmogorov and the Vitushkin results. In practical examples, evaluation of the  $\nu$ -jets has to be based on input-values clustered around the data base  $D$ .



We can now specify the assumptions of our main result, theorem 1.1 below, on the identification of scalar hybrid models. On the tree  $\Gamma$  we assume monotonicity:

$$u_{j,x_i}(x^j) \neq 0 \quad (1.10)$$

for all partial derivatives  $x_i$ ,  $i \in p_j$ , at any black or white box  $j \in W \cup B$ , and at any point  $x^j = (x_i)_{i \in p_j} \in \mathbb{R}^{|p_j|}$ . The nonmonotone case in fact poses serious obstacles to unique identification of black boxes, which we briefly discuss in section 7; see (7.2), (7.3) below.

Our second assumption concerns the white-box functions  $u_j$ ,  $j \in \Gamma'$  on their input data  $D_j := \{x^j = (x_i)_{i \in p_j}; x^0 \in D\}$ . We assume

$$\text{genericity of } u_j|_{D_j}, \text{ for all } j \in W \quad (1.11)$$

in the  $C^\nu$  topology. Here and below, the word *genericity* indicates that the corresponding assumption, or statement, is assumed, or claimed, to hold for a countable intersection of open-dense sets in the  $C^\nu$  Whitney topology, see [14].

The monotonicity assumption (1.10) restricts our considerations to an open subset of  $C^\nu$ , in this topology. We then still talk about genericity. Obviously, density will only be required to hold within the subset (1.10).

With simple addition as a white box, for example, it is obvious that the three layer tree

$$u_1(x_1) + u_2(x_2) = h(x_1, x_2) \quad (1.12)$$

of Figure 1.1 provides a counterexample to unique identification of the black boxes  $u_1$  and  $u_2$ . Indeed, with any solution pair  $u_1, u_2$  and any real constant  $c$ , we also have solution pairs

$$u_1(x_1) + c, \quad u_2(x_2) - c. \quad (1.13)$$

Our genericity assumption (1.11) on white boxes avoids this source of nonuniqueness.

Subsequent black boxes  $j$ ,  $j' = s(j)$ , in contrast, are an unavoidable structural source of nonuniqueness. Indeed, let  $u_j$  and  $u_{j'}$  be any solution pair and choose any real diffeomorphism  $\sigma \in C^\nu(\mathbb{R}, \mathbb{R})$ . Then we also have a solution pair

$$u_{j'} \circ \sigma, \quad \sigma^{-1} \circ u_j. \quad (1.14)$$

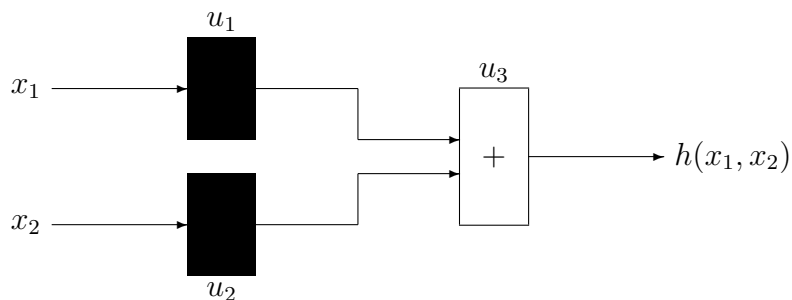


Figure 1.1: Nonunique identification due to additive white box.

Here  $\sigma$  preceding  $u_{j'}$  is understood to act on the  $x_j$ -entry of the inputs to  $u_{j'}$  only. We call the universal, unavoidable construction (1.14) a *trivial calibration* by  $\sigma$ . The calibration (1.13) of a shift by  $\pm c$ , in contrast, is not universal but depends on the particular additive structure of the white box  $u_3$  in Figure 1.1. It disappears, in fact, under small perturbations of  $u_3$ . Moreover, the black boxes  $u_1, u_2$  in Figure 1.1 are neighboring, and not subsequent to each other.

In a similar spirit, we may view white boxes with only one single input as simple coordinate transformations of the output of the preceding white or black box. We will discard this trivial case and assume below that any white box possesses at least two inputs.

Our final assumptions concern the connected  $d$ -dimensional data base  $D \subset \mathbb{R}^{|\ell_0|}$  of inputs  $x^0$  on which the  $\nu$ -jet of the global input-output data  $h(x^0)$  is assumed to be known. To this end, we introduce the notion of a *transverse data base*. We consider any input  $x_i$ ,  $i \in p_j$ , to any black or white box  $u_j$ . Note how  $x_i$  depends only on variables  $x_\iota$  for some index subset  $\iota \in J_i \subset \ell_0$  of the input layer. By the tree structure of our hybrid model, the sets  $J_i$  are disjoint, for different  $i$  belonging to the same layer  $\ell_k$ , and these sets  $J_i$  decompose the input layer  $\ell_0$ . By monotonicity assumption (1.10) and the chain rule, the partial derivatives of  $x_i$  with respect to any input  $x_\iota$ ,  $\iota \in J_i$ , are nonzero. Because the sets  $J_i$  are disjoint, in particular, for inputs  $i \in p_j$  of any box  $j$ , the function  $(x_i)_{i \in p_j}$  of the global input vector  $x^0 \in \mathbb{R}^m$  possesses only regular values. For every black box  $j \in B$ , the values of its input vector  $x^j = (x_i)_{i \in p_j} \in D_j$  therefore define a foliation of the input space  $\mathbb{R}^{|\ell_0|}$  with leaves of codimension  $|p_j|$ . We call the data base  $D$  *minimally transverse* if  $D$

possesses dimension

$$d := \dim D = \max_{j \in B} |p_j| \quad (1.15)$$

and is transverse to these foliations, for any black box  $j \in B$ . Moreover we assume the level sets

$$D_{j,0}^c := \{x^0 \in D; u_j(x^j(x^0)) = c\} \quad (1.16)$$

to be (empty or) *connected* in the data base  $D$ , for all real values  $c$  and all black boxes  $j \in B$ . Here and below it is convenient to let  $x^j$  also denote the resulting map from the global input vector  $x^0$  to the input vector  $x^j = (x_i)_{i \in p_j}$  of box  $j$ .

As a consequence of the transversality assumption on the  $d$ -dimensional data base manifold  $D$ , the functions  $x^j = (x_i)_{i \in p_j}$  possess only regular values on  $D$ , in fact with surjective linearization, for any black box  $j \in B$ . Moreover each input set  $D_j = x^j(D) \subseteq \mathbb{R}^{p_j}$  is then open and, as  $D$  itself, connected.

We illustrate the above assumption by a simple example. Since transversality is based on linear approximation, we choose a completely linear example, for global simplicity. We thus represent all boxes  $u_j = u_j(x^j) = \tilde{u}_j^T x^j$  by row vectors  $\tilde{u}_j^T \in (\mathbb{R}^{|p_j|})^*$  with positive components. Similarly, the resulting value  $x_j \in \mathbb{R}^{|p_j|}$  can be represented by a matrix  $x^j(x^0) = X^j x^0$  with  $|p_j|$  rows and  $|\ell_0|$  columns. The entries in rows  $i \in p_j$  and associated columns  $\iota \in J_i$  of the matrix  $X^j$  are strictly positive; all other entries are zero. Because the sets  $J_i$  are disjoint, the matrix  $X^j$  is surjective. We now choose the data base  $D$  to be affine linear:  $D = c^0 + \ker \tilde{D}$  for some full rank matrix  $\tilde{D}$  of  $|\ell_0|$  columns and  $|\ell_0| - d$  rows. Clearly  $d = \dim D$ .

Fix any value  $c^j \in (\mathbb{R}^{|p_j|})$  for  $x^j$ . Then transversality of the data base  $D$  to the leaf  $X^j x^0 = c^j$  of inputs  $x^0$  holds, if, and only if, the matrix

$$\begin{pmatrix} X^j \\ \tilde{D} \end{pmatrix} : \mathbb{R}^{|\ell_0|} \rightarrow \mathbb{R}^{|p_j|+|\ell_0|-d} \quad (1.17)$$

is surjective. Because  $d \geq |p_j|$  and  $X^j : \mathbb{R}^{|\ell_0|} \rightarrow \mathbb{R}^{|p_j|}$  is surjective, this holds automatically, for almost all choices of the matrix  $\tilde{D}$ , both in the Lebesgue and in the open-dense sense. Connectedness of the linear spaces  $D_{j,0}^c$  is trivially satisfied. Our above transversality and connectedness assumptions (1.15), (1.16) are nonlinear generalizations of this simple illustration.

After unique identification of all black boxes, the resulting identified structured hybrid model will allow for extrapolation of the global input-output data  $h(x^0)$ , possibly far beyond its original domain  $x^0 \in D$ . Using the input vectors  $x^j = x^j(x^0) = (x_i)_{i \in p_j}$  of black boxes  $j \in B$ , as introduced in (1.6), (1.7), we define the *spanned set*  $\bar{D}$  of inputs  $x^0$  as

$$\begin{aligned} \bar{D} &:= \{x^0 \in \mathbb{R}^{|\ell_0|}; x^j \in D_j \text{ for all } j \in B\} = \\ &= \bigcap_{j \in B} (x^j)^{-1}(D_j) = \bigcap_{j \in B} (x^j)^{-1}(x^j(D)). \end{aligned} \quad (1.18)$$

For example, if black boxes  $j \in B$  arise only in the first layer,  $B \subset \ell_1$ , then  $D_j$  is simply the coordinate projection of the data base  $D$  to the input vector  $x^j$  of the black box  $j$ . The spanned set  $\bar{D}$  of inputs is then given by

$$\bar{D} = \bigotimes_{j \in B} D_j \times \bigotimes_{j \in W \cap \ell_1} \mathbb{R}^{|p_j|} \quad (1.19)$$

and hence possesses full dimension

$$\dim \bar{D} = \sum_{j \in B} |p_j| + \sum_{j \in W \cap \ell_1} |p_j| = |\ell_0|. \quad (1.20)$$

Of course this total dimension  $|\ell_0|$  of the global input vector  $x^0$  will in general much exceed the dimension  $d = \max_{j \in B} |p_j|$  of our minimally transverse data base  $D$ . For another example, with  $d = 1$  and arbitrary dimension  $|\ell_0| = \dim \bar{D} \geq n$ , we refer to the black-white-black trees discussed in section 3 below; see (3.55).

We are now ready to state our main result. For a previous presentation without proof see [18].

**Theorem 1.1** *Consider a structured hybrid model (1.9) on a tree  $\Gamma$ . Let assumptions (1.10), (1.11), (1.15), (1.16) hold for the global input-output function  $h$  on the minimally transverse data base  $D$ , and for the scalar white-box and black-box functions  $u_j$ , all of class  $C^\nu$  with  $\nu \geq 5$ . We also assume that any white box has at least two inputs.*

*Then the black-box functions  $u_j$ ,  $j \in B$ , are determined uniquely on their respective input data  $x^j = (x_i)_{i \in p_j}$  in  $D_j$ , up to trivial calibrations of the form (1.14).*

*In particular the global input-output function  $h$  can be extrapolated uniquely from the lower-dimensional data base  $D$  to the full spanned set  $\bar{D}$  of inputs  $x^0 = (x_j)_{j \in \ell_0}$  defined in (1.17) above.*

We emphasize the recursive aspect of this theorem. For black boxes  $j$  in the first layer  $\ell_1$  of the tree,  $j \in B \cap \ell_1$ , the input data set  $D_j$  is simply the projection of the minimally transverse data base  $D$  onto the vector component  $x^j$  of the global input vector  $x^0 \in \mathbb{R}^{|\ell_0|}$ . With these input sets defined, a priori, all black boxes  $j \in B \cap \ell_1$  of the first layer are determined uniquely, up to trivial calibration. Together with the remaining, white boxes  $j \in W \cap \ell_1$  this defines the input data sets  $D_j$  uniquely at the second layer  $j \in \ell_2$ , and so on through all layers  $\ell_3, \ell_4, \dots, \ell_m$  of the tree  $\Gamma$ . Since our proof in section 5 will proceed constructively along the same recursion, it is in fact possible to check the monotonicity, genericity, and transversality assumptions (1.10), (1.11), (1.15), (1.16) successively, in a practical implementation.

The remaining sections are organized as follows. In section 2 we show how a lemma of Frobenius type reduces the identification problem to only one elementary type of a three layer hybrid model: the *black-white-black tree*. In section 3 we describe the unique identification of this remaining elementary type, by local normal forms. Section 4 is devoted to the genericity assumptions on  $u_j$ . For a summary of the proof of theorem 1.1 see section 5. After a specific polymerization example from chemical industry, presented in section 6, we conclude with a summary and discussion, in section 7.

**Acknowledgment.** This work was supported by the Bundesministerium für Bildung und Forschung of the Federal Republic of Germany under the projects “Identifikation verketteter Reaktoren mit Methoden der nichtlinearen Dynamik” and “Verfahren zur Identifikation hybrider, verfahrenstechnischer Prozessmodelle”. We are much indebted to Vassili Gelfreich, Marc Georgi, Nihar Jangle, Tobias Marxen, Stefan Liebscher, Alexander Mielke, and Dima Turaev, for helpful discussions. The referees are to be praised for constructively helping to clarify the precise role of the data base in our analysis.

## 2 Reduction to the elementary black-white-black tree

In this section we reduce the proof of our main theorem 1.1 to a detailed analysis of only one single type of an elementary three layer tree: a black-white-black tree with intermediate white box. See Figure 2.1 below. The fundamental step in this reduction will be an identification of the level sets of black-box functions  $u_j$  via a Frobenius-like theorem given in lemma 2.1.

For alternating 1-forms  $\omega$ , the Frobenius theorem asserts the existence of manifolds of codimension 1 such that  $\omega$  vanishes on their tangent spaces. Such manifolds exist if, and only if,  $\omega \wedge d\omega \equiv 0$ . Below we consider the simple case that  $\omega = du$  is a gradient. For the convenience of the reader we include an elementary proof.

**Lemma 2.1** *Let  $\mathcal{D} \subset \mathbb{R}^n$  be open and connected. Let  $u, \tilde{u} \in C^\nu(\mathcal{D}, \mathbb{R})$  possess nonvanishing parallel gradients:*

$$u_\xi(\xi) = \alpha(\xi)\tilde{u}_\xi(\xi) \neq 0 \quad (2.1)$$

*for some positive scalar function  $\alpha \in C^{\nu-1}(\mathcal{D}, \mathbb{R})$ . In addition assume the level sets of  $u$  and  $\tilde{u}$  to be connected. Then*

$$u = \sigma \circ \tilde{u}, \quad (2.2)$$

*for some scalar function  $\sigma \in C^\nu(\mathbb{R}, \mathbb{R})$  with positive derivative  $\sigma' > 0$ .*

*An analogous statement holds for negative  $\alpha$  and  $\sigma'$ .*

**Proof:** Consider any  $C^\nu$ -curve  $\xi(s) \in \mathcal{D}$  in any level set  $\tilde{u} \equiv \text{const}$ . Then

$$\frac{d}{ds}u(\xi(s)) = u_\xi \dot{\xi} = \alpha \tilde{u}_\xi \dot{\xi} = \alpha \frac{d}{ds}\tilde{u}(\xi(s)) = 0. \quad (2.3)$$

Hence  $\xi(s)$  also lies in a single level set of  $u$ . Since level sets are connected, by assumption, level sets of  $\tilde{u}$  are therefore contained in level sets of  $u$ . Similarly, interchanging the roles of  $u$  and  $\tilde{u}$ , level sets of  $u$  are contained in level sets of  $\tilde{u}$ . Since the level sets of  $u$  and  $\tilde{u}$  thus coincide, we see that (2.2) holds globally. Choosing  $\xi(s)$  transverse to the level sets, in contrast, we conclude  $\sigma \in C^\nu$  with

$$\sigma'(\tilde{u}(\xi(s))) = \alpha(\xi(s)). \quad (2.4)$$

This proves the lemma. ⊗

In our proof of theorem 1.1 we plan to apply this lemma to the unique identification of black box functions  $u_j$ ,  $j \in B$ , on their open and connected input data sets  $x^j \in D_j \subseteq \mathbb{R}^{|p_j|}$ . We therefore show next that the level sets

$$D_j^c := \{x^j \in D_j; \quad u_j(x^j) = c\} \quad (2.5)$$

are indeed connected.

**Lemma 2.2** *Let the data base  $D$  be minimally transverse, see (1.15), with connected level sets  $D_{j,0}^c$ , see (1.16). Then for any black box function  $u_j$ ,  $j \in B$ , satisfying monotonicity assumption (1.10), and for any real value  $c$ , the level set  $D_j^c$  is empty, or else is a connected submanifold of the open input set  $D_j$ , with codimension one.*

**Proof:** Because  $D$  is assumed to be transverse, and because  $u_j$  satisfies monotonicity assumption (1.10), any value  $c$  is a regular value of  $u_j \circ x^j$  on  $D$ . In particular the level sets

$$D_{j,0}^c := \{x^0 \in D; \quad u_j(x^j(x^0)) = c\}, \quad (2.6)$$

if nonempty, are codimension-1 submanifolds of  $D$  and foliate  $D$ . Note that the range  $I$  of the continuous function  $u_j \circ x^j$  on  $D$  is connected, because  $D$  is connected, and hence  $I$  is an (open) interval. For  $c \in I$ , each level set  $D_{j,0}^c$  is nonempty and, by assumption (1.16), also connected. In particular, the continuous image

$$D_j^c = x^j(D_{j,0}^c) \quad (2.7)$$

is also connected, and the lemma is proved. ⊗

We now apply the lemmata to any black box  $u = u_j(x_i)_{i \in p_j}$  of the first layer,  $j \in \ell_1$ , with input vector  $\xi := x^j = (x_i)_{i \in p_j} \in D_j \subseteq \mathbb{R}^{|p_j|}$ . For a suitable scalar-valued  $C^\nu$ -function  $\Psi$  which encodes the remaining hybrid model, and with a canonical decomposition

$$x^0 = (\xi, \eta) \in D \subseteq \mathbb{R}^{|\ell_0|} \quad (2.8)$$

we can then write

$$\Psi(u(\xi), \eta) = h(\xi, \eta). \quad (2.9)$$

Differentiation with respect to  $\xi$  yields

$$\Psi_u \cdot u_\xi = h_\xi. \quad (2.10)$$

Now suppose  $u, \tilde{u}$  are two solutions of the hybrid model, at box  $j$ . Then  $u_\xi(\xi) = \alpha(\xi)\tilde{u}_\xi(\xi)$  for some nonzero scalar function  $\alpha \in C^{\nu-1}$ , as required by assumption (2.1) of lemma 2.1. Indeed, the transversality assumption (1.15) on the data base  $D$  allows us to invert the first component  $s \mapsto \xi(s)$  of its  $C^\nu$ -parametrization  $s \mapsto (\xi(s), \eta(s))$  along a suitable  $|p_j|$ -dimensional submanifold of the database  $D$ . As a consequence,

$$\alpha(\xi) := \Psi_u(\tilde{u}(\xi), \eta) / \Psi_u(u(\xi), \eta) \quad (2.11)$$

can be written as a function of  $\xi$  alone. Therefore lemma 2.1 implies

$$u_j = \sigma_j \circ \tilde{u}_j, \quad (2.12)$$

i.e., uniqueness of the black box  $u_j$  up to a calibration of its values by a scalar  $C^\nu$ -function  $\sigma_j$ . Indeed, connectedness of the level sets of  $u, \tilde{u}$ , as required by lemma 2.1, has been established in lemma 2.2.

In passing, we note that the level surfaces of  $u_j$  can be determined constructively, from the gradients (2.10), by successive solution of families of suitable ordinary differential equations; see [3].

The successor  $j' := s(j)$  of the black box  $j$ , can either be black again,  $j' \in B$ , or else white,  $j' \in W$ . If  $j' \in B$ , then we may replace  $u_j$  by any other solution  $\tilde{u}_j = \sigma_j^{-1} \circ u_j$ , and absorb  $\sigma_j^{-1}$  into the black successor box  $u_{j'}$  by trivial calibration; see (1.14). Thus  $u_j$  has been identified uniquely, up to an unavoidable trivial calibration.

Next suppose the successor box is white,  $j' \in W$ . Analogously to (2.12), and up to calibration by scalar functions  $\sigma_j$ , we eliminate all other black boxes  $x_j = \sigma_j(\tilde{u}_j((x_i)_{i \in p_j}))$ ,  $j \in p_{j'}$ , which also provide inputs to the same white box  $j'$ . Note that the calibrations  $\sigma_j$  cannot be absorbed into the successor box  $u_{j'}$  by trivial calibration, because  $j' \in W$  is a white box this time. We thus arrive at a new hybrid model of the form

$$\Phi(u_{j'}((\sigma_j(x_j))_{j \in p_{j'}}), \eta) = h((x_j)_{j \in p_{j'}}, \eta). \quad (2.13)$$



Here we consider the variables  $x_j = \tilde{u}_j((x_i)_{i \in p_j})$  as new inputs, because they are preceded by only white boxes  $\tilde{u}_j$ , and thus are uniquely determined by the original input vector  $x^0$ . As above,  $\eta$  denotes the remaining input components.

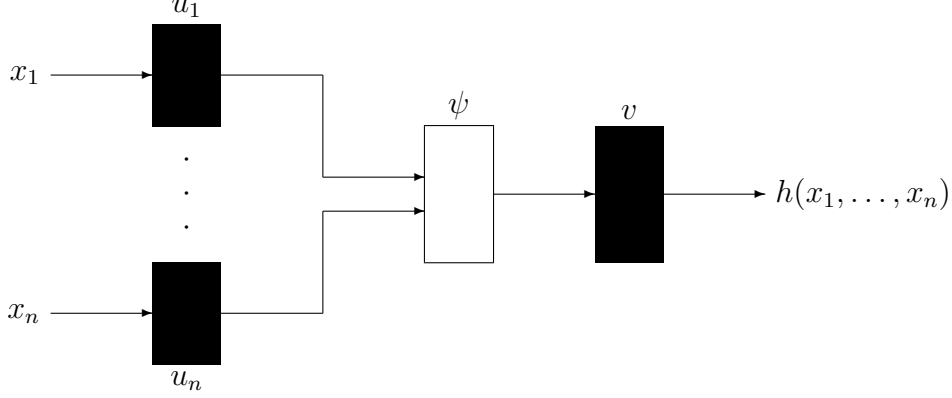


Figure 2.1: The elementary black-white-black (bwb) hybrid tree.

We now view  $\eta = \eta_0$  in (2.13) as a constant, and define the scalar function

$$v := \Phi(\cdot, \eta_0) \in C^\nu, \quad (2.14)$$

which we consider as a new black box. With the abbreviation  $h_0 := h(\cdot, \eta_0)$  we can then rewrite (2.13) as

$$v(u_{j'}((\sigma_j)_{j \in p_{j'}})) = h_0. \quad (2.15)$$

See Figure 2.1 for an illustration of this elementary *black-white-black hybrid tree*. Note the slightly different notation  $\sigma_j \mapsto u_j$ ,  $u_{j'} \mapsto \psi$ ,  $h_0 \mapsto h$ , which distinguishes the only white box  $u_{j'}$  by the Greek letter  $\psi$ , and reminds of the origin of the calibrations  $\sigma_j$  from the black boxes  $u_j$ .

Unfortunately, the  $d$ -dimensional data base  $D$  does not allow us to vary the inputs  $x_j$ , independently, and at the same time freeze the remaining input variables at values  $\eta_0$ . As far as partial derivatives with respect to  $x_j$  at a fixed value  $x_j = x_j^0$  are concerned, for  $j \in p_{j'}$  and up to any finite order, however, we may well consider (2.15) as valid. In sections 3 and 4 we will therefore show how to uniquely determine all black boxes  $u_j$ , viz. the above calibrations  $\sigma_j$ , in this bwb-tree, from local information on the  $\nu$ -jet of  $\tilde{h} \in C^\nu$ . This will complete the identification of all black boxes  $u_j$  – unique up to trivial calibration – in the original first layer  $j \in \ell_1$  of the hybrid tree model.

By induction on the total number  $m$  of layers, this will eventually allow us to also complete the identification of the entire tree.

### 3 Identification of elementary black-white-black trees by local normal forms

In this section, we solve the identification problem of the scalar black-white-black tree

$$v(\psi(u_1(x_1), \dots, u_n(x_n))) = h(x_1, \dots, x_n) \in \mathbb{R}, \quad (3.1)$$

which was posed in (2.15); see also Figure 2.1. We develop a local normal form approach which, for generic white-box functions  $\psi$ , allows the unique global identification of the black-box functions  $u_j$  and  $v$ , from a one-dimensional data base  $D$ . In particular, we begin to prove theorem 1.1 for the special case of the structured hybrid model (3.1), which does not allow for any remaining trivial calibrations.

Normal forms are representatives of group orbits. The Jordan normal form of linear algebra, for example, represents orbits of square matrices  $A$  under conjugation  $SAS^{-1}$  by the group of invertible matrices  $S$ . The Poincaré-Birkhoff normal form represents jets of vector fields under transformation by local diffeomorphisms; see e.g. [21] and the references there. From this abstract view point, equation (3.1) states that  $h$  and  $\psi$  lie in the same group orbit under the action of transformations by diffeomorphisms  $u_1, \dots, u_n, v$  of the real line. For simplicity, we consider smooth functions; the obvious adaptation to  $C^\nu$ -functions just limits considerations to using at most  $\nu$  derivatives.

The local study of smooth functions, like  $\psi$ , and of their Taylor expansions under local diffeomorphisms in their source, here  $\mathbb{R}^n$ , as well as their target, here  $\mathbb{R}$ , is the subject of singularity theory; see for example [4]. The arising normal forms are sometimes also studied under further constraints on the full group of diffeomorphisms. Dufour, for example, considers collections of functions with shared sources and targets, which have to be transformed in parallel, by the same diffeomorphisms [9]. His setting is different from ours, even though his results are expressed in a graph language. For our specific and rather restricted identification problem, in the

present paper, we find it most convenient – and more explicit – to proceed directly, and on an elementary level, with our local analysis.

**Lemma 3.1** *Let  $\psi$  satisfy  $\psi(0) = 0$  with the following nondegeneracy condition, at  $u = 0$  :*

$$\psi_{u_j}(0) \neq 0. \quad (3.2)$$

*Then there exist locally diffeomorphic transformations  $u_1, \dots, u_n, v$  such that*

$$v(\psi(u_1(x_1), \dots, u_n(x_n))) = H(x_1, \dots, x_n), \quad (3.3)$$

*where the normal form  $H$  of  $\psi$  satisfies*

$$H(t\mathbf{e}_j) = t, \quad (3.4)$$

$$H(t\mathbf{e}) = nt. \quad (3.5)$$

*Here  $\mathbf{e}_j$  denotes the  $j$ -th unit vector in  $\mathbb{R}^n$  and  $\mathbf{e} = (1, \dots, 1)$ .*

*Under the normalizations*

$$u_j(0) = 0, \quad v(0) = 0, \quad v'(0) = 1, \quad (3.6)$$

*the jets of the transformations  $u_j, v$  at zero are uniquely determined in dimensions  $n \geq 2$ .*

**Proof:** To prove existence of the normal form transformations, we only consider the nontrivial case  $n \geq 2$ . We use the abbreviations  $x = (x_1, \dots, x_n)$ ,  $u(x) = (u_1(x_1), \dots, u_n(x_n))$  and

$$\psi_j(t) := \psi(t\mathbf{e}_j); \quad \varphi(t) := \psi(\psi_1^{-1}(t), \dots, \psi_n^{-1}(t)). \quad (3.7)$$

We also use and identify the inverse  $v^{-1}$  of  $v$ , instead of  $v$  itself. Inserting  $x = t\mathbf{e}_j$  in the resulting variant

$$\psi(u(x)) = v(H(x)) \quad (3.8)$$

of (3.3), we see that the normal form condition (3.4) is equivalent to

$$\psi_j(u_j(t)) = \psi(u_j(t)\mathbf{e}_j) = \psi(u(t\mathbf{e}_j)) = v(H(t\mathbf{e}_j)) = v(t). \quad (3.9)$$

Here we have used the normalizations  $u_j(0) = 0$  of (3.6). Clearly (3.9) holds if, and only if, we choose

$$u_j := \psi_j^{-1}v. \quad (3.10)$$

Inserting  $x = t\mathbf{e}$  into (3.8), next, we see that the normal form condition (3.5) is equivalent to

$$\begin{aligned}\varphi(v(t)) &= \psi(\psi_1^{-1}(v(t)), \dots, \psi_n^{-1}(v(t))) = \psi(u_1(t), \dots, u_n(t)) = \\ &= \psi(u(t\mathbf{e})) = v(H(t\mathbf{e})) = v(nt).\end{aligned}\tag{3.11}$$

Since  $\varphi(0) = 0$ ,  $\varphi'(0) = n$ , we see that (3.11) amounts to a conjugation of  $\varphi$  to its linearization, by  $v$ . By standard normal form theory for diffeomorphisms  $\varphi$ , such a conjugation  $v$  satisfying the normalization (3.6) is indeed possible for  $n \neq 0, \pm 1$ . (For an explicit formula for the transformation  $v(t)$  see (4.6) below.) In particular, this proves the existence of a normal form (3.3)–(3.5), for dimension  $n \geq 2$ . We remark that the derivatives  $v^{(k)}(0)$ ,  $k \geq 2$ , can be obtained, recursively, by  $k - 1$  differentiations of  $\varphi'(v(t)) \cdot v'(t) = nv'(nt)$  at  $t = 0$ :

$$nv^{(k)}(0) + \dots = \varphi'(0)v^{(k)}(0) + \dots = n^k v^{(k)}(0),\tag{3.12}$$

where the omitted terms contain derivatives of  $v$  of order at most  $k - 1$ .

To prove uniqueness of the jets, under the normalization (3.6), we first invoke the group property of the local transformations. It is therefore sufficient to show that

$$H(u(x)) = v(H(x))\tag{3.13}$$

for  $H$  in normal form (3.4), (3.5) implies  $J(u_j) = J(v) = J(\text{id})$  for the  $k$ -jets  $J$  at 0. Here we have substituted  $H$  itself for  $\psi$  in (3.8). This reduces our previous analysis to the special case  $\psi_j = \text{id}$ ,  $\varphi = n \cdot \text{id}$ , by (3.4), (3.5). Therefore  $u_j \equiv v$ , by (3.10) and  $nv(t) = v(nt)$ , by (3.11). Differentiating  $k$  times at  $t = 0$ , we obtain

$$nv^{(k)}(0) = n^k v^{(k)}(0).\tag{3.14}$$

The normalization (3.6) implies  $v(0) = 0$  and  $v'(0) = 1$ . The  $k$ -jets of  $v$  and  $\text{id}$  therefore coincide at  $t = 0$ . Since  $u_j \equiv v$ , the same holds true for all transformations  $u_j$ . This proves the lemma.  $\boxtimes$

We now provide an explicit condition on the normal form  $H$  of (3.4), (3.5), under which jet uniqueness of the normal form transformations  $u_j, v$  remains valid, even without the somewhat artificial constraint  $v'(0) = 1$ . The one remaining free choice of  $v'(0)$  will be used to further restrict the normal form.

**Corollary 3.2** *Assume the normal form  $H$  of lemma 3.1 satisfies the additional nondegeneracy condition*

$$H_{k_0} \neq 0 \tag{3.15}$$

*for some nonzero derivative multi-index  $k_0$  of even order  $|k_0|$ . Then jet uniqueness at zero holds for the transformations  $u_j, v$ , if we augment the normal form conditions (3.5), (3.6) by the additional restriction*

$$H_{k_0} = 1. \tag{3.16}$$

*Here we still require the transformations  $u_j, v$  to satisfy*

$$u_j(0) = 0, \quad v(0) = 0. \tag{3.17}$$

**Proof:** Without loss of generality, we may assume  $\psi = H$  to be in normal form (3.4), (3.5) already. Linear transformations  $v^{-1} \circ \psi \circ u$  with  $v(t) = u_j(t) = \lambda t$  and nonzero  $\lambda$  do not affect this fact. For the  $k$ -th derivative at  $x = 0$ , however, they imply

$$(v^{-1}\psi u)^{(k)}(0) = \lambda^{|k|-1}\psi^{(k)}(0). \tag{3.18}$$

For  $k = k_0$  of even order  $|k_0|$ , this allows us a normalization of the nonzero derivative  $H_{k_0}$  of the normal form, for example as specified in (3.15), (3.16).

To prove jet uniqueness of the normalizing transformations  $u_j, v$ , we argue as in (3.13), (3.14) above. Clearly,  $v$  possesses a purely linear jet at zero. We denote this fact by writing  $v(t) = \lambda t$ , for some nonzero  $\lambda$ . Similarly  $u_j(t) = \lambda t$  in (3.13), by (3.10). Therefore the jet of  $H$  at  $x = 0$  satisfies  $H(\lambda x) = \lambda H(x)$ , by (3.3). More precisely

$$\lambda^{|k|}H^{(k)}(0) = \lambda H^{(k)}(0), \tag{3.19}$$

for all  $k = (k_1, \dots, k_n)$ . For  $k = k_0$  of even order  $|k_0|$ , we see that the additional normalization (3.16) implies  $v'(0) = \lambda = 1$ . This proves the corollary.  $\boxtimes$

We observe that other normalization conditions than (3.16) are feasible. For example, the specified even order derivative  $k_0$  can be discarded in favor of any nontrivial linear functional of partial derivatives of fixed even order  $|k| \neq 0$ . Odd order derivatives  $|k_0| \neq 1$  only determine  $v'(0) = \pm 1$ , up to a sign.

p \ n	2	3	4	5
0	0	0	0	0
1	0	0	0	0
2	0	1	4	8
3	1	7	19	37
4	2	18	49	101
5	5	35	100	221

Table 3.1: Dimensions  $d(n, p)$  of spaces of Taylor coefficients  $H_k$  with  $|k| \leq p$  satisfying the normal form constraints (3.20)–(3.23).

Because of its central importance for our solution of the identification problem of black-white-black trees, and of scalar hybrid tree models in general, we briefly summarize the properties of the jet of the normal form  $H$ .

**Corollary 3.3** *For multi-indices  $k = (k_1, \dots, k_n)$  let  $H_k := H^{(k)}(0)$  denote the corresponding partial derivative of the normal form  $H$  at zero. Then the normal form conditions (3.4), (3.5), (3.16) imply*

$$H_k = 0, \quad \text{for } k = p \cdot e_j \quad (3.20)$$

$$H_k = 1, \quad \text{for } k = e_j \quad (3.21)$$

$$H_{k_0} = 1 \quad (3.22)$$

$$\sum_{|k|=p} \frac{1}{k!} H_k = 0 \quad (3.23)$$

hold for all orders  $p \neq 1$  and for all  $j = 1, \dots, n$ . The even order  $|k_0|$  can be chosen to be 4, for  $n = 2$ . For  $n \geq 3$ , we can choose  $|k_0| = 2$ . Of course  $k_0 \neq p e_j$  must not contradict (3.20).

Conversely, analytic functions  $H$  with Taylor coefficients satisfying (3.20) – (3.23) are in normal form (3.4), (3.5), (3.16).

**Proof:** The proof is immediate from conditions (3.4), (3.5), (3.16) on the normal form  $H$ . ⊠

**Corollary 3.4** *The dimension  $d(n, p)$  of the linear space of real Taylor coefficients  $H_k$  of order  $|k| \leq p$ , satisfying the normal form constraints (3.20)–(3.23), is given by*

$$d(n, p) = \begin{cases} 0, & \text{for } p = 0, 1 \\ \binom{n+p}{n} - (n+1)p - 1, & \text{for } p \geq 2 \end{cases} \quad (3.24)$$

for  $n \geq 3$ . For  $n = 2$  and  $p = 2, 3$ , the numbers in (3.24) have to be increased by 1, due to  $|k_0| = 4$ . For the resulting dimensions  $d(n, p)$  with  $p, n \leq 5$  see Table 3.1.

**Proof:** The proof follows by elementary counting.  $\boxtimes$

**Corollary 3.5** *Lemma 3.1 and corollaries 3.2–3.4 remain valid, if we replace the reference point  $u = 0$  by  $u = u^0$ . Specifically, we then assume*

$$\psi_{u_j}(u^0) \neq 0. \quad (3.25)$$

*Jet uniqueness of the transformations  $u_j, v$ , in particular, remains valid with the normalizations*

$$u_j(0) = u_j^0, \quad v(\psi(u^0)) = 0, \quad (3.26)$$

*and, in lemma 3.1, with the additional requirement  $v'(\psi(u^0)) = 1$ . We note that  $v(\psi(u^0)) = 0$  already follows from (3.3).*

Let  $I = I(H, 0)$  be any function which depends on the  $\nu$ -jet of the normal form  $H$ , evaluated at  $x = 0$ . More generally, define,

$$I(\psi, u^0) := I(H, 0) \quad (3.27)$$

for functions  $\psi$  which lie in the group orbit of  $H$  under the normal form transformation (3.3). Note that  $I(H, 0)$  is well-defined, if we assume nondegeneracy condition (3.15) on  $H$ . We have normalized  $H$  by (3.4), (3.5), (3.16); see corollaries 3.2, 3.3.

We now assume (3.1) to hold:

$$v(\psi(u(x))) = h(x). \quad (3.28)$$

In other words, the given white-box function  $\psi$  and the global input-output data  $h$  lie in the same group orbit under the action of the transformations  $u, v$ . In particular, they possess the same normal form  $H$ . Therefore

$$I(\psi, u^0) = I(H, 0) = I(h, x^0). \quad (3.29)$$

We have thus proved the following theorem, fundamental to the identification of black-white-black trees and, more generally, of scalar hybrid models of tree structure.

**Theorem 3.6** *Let  $I(H, 0)$  be any function of the  $\nu$ -jet of nondegenerate normal forms  $H$ , as defined above. Let (3.1) hold,  $v(\psi(u(x))) = h(x)$ , with black- and white-box functions  $u_j$ ,  $v$ , and  $\psi$ , and with global input-output data  $h$ ; all with nonvanishing first (partial) derivatives. Then*

$$I(\psi, u^0) = I(h, x^0) \quad (3.30)$$

holds at  $u^0 := u(x^0) = (u_1(x^0), \dots, u_n(x^0))$ , where  $I(\psi, u^0)$  and  $I(h, x^0)$  are both defined as in (3.27).

Note that  $I(\psi, u^0)$ ,  $I(h, x^0)$  can be expressed by partial derivatives of the given functions  $\psi$ ,  $h$  at  $u^0$ ,  $x^0$ , respectively. We can therefore expect (3.30), typically, to be solvable for  $u^0$  as a function of  $x^0$ :

$$u^0 = u(x^0). \quad (3.31)$$

This solution determines the black-box functions  $u_j$  at  $x^0$ . By  $v(\psi(u)) = h(x)$ , it also determines the black-box function  $v$  at  $\psi(v(x^0))$ . Because  $x^0$  was arbitrary, this completes the identification process. The unique solvability of systems (3.30) by (3.31) will be addressed in section 4.

Theorem 3.6 may, by now, seem a mathematical tautology to some abstract-minded reader. Indeed it only summarizes our normal form approach. In conjunction with our previous explicit normal form derivation, however, the theorem yields highly nontrivial and useful explicit expressions for the practical identification of all black boxes  $u_j$  and  $v$ . Because this is the central building block to our identification procedure, we conclude this section with a complete set of explicit expressions for identifying equations of bwb trees, for all dimensions  $n \geq 2$ . In reminiscence of our introductory remarks on Vitushkin complexity, and in view of Table 3.1, we also illustrate how higher derivatives of the data  $h$  and  $\psi$  are required for identification of systems with a smaller number  $n$  of input variables.

Here are our examples. We first compute the group invariants

$$I_k(\psi, u^0) = H_k \quad (3.32)$$



for  $n \geq 4$  inputs and order  $p = |k| \leq 2$ . Following corollary 3.3, we consider the second order normalization

$$H_{x_1 x_2} = 1. \quad (3.33)$$

With the normal form (3.4), (3.5), (3.32) we thus expect

$$d(n, p) = \frac{1}{2}n(n-1) - 2 \quad (3.34)$$

invariants, in terms of the Taylor coefficients

$$\begin{aligned} \psi_j &:= \psi_{u_j}(u^0), \\ \psi_{j_1 j_2} &:= \psi_{u_{j_1} u_{j_2}}(u^0) \end{aligned} \quad (3.35)$$

of  $\psi$  at  $u = u^0$ , of order  $|k| \leq p = 2$ .

We first show how the normal form transformations

$$v(\psi(u(x))) = H(x) \quad (3.36)$$

are determined, up to second order, by the normal form constraints (3.20), (3.21), (3.23), (3.33). Instead of following the reduction of lemma 3.1 and its corollaries, we proceed by brute force. Differentiating (3.36) with respect to  $x_j$ , at  $x = 0$ , we obtain

$$1 = H_{x_j} = v' \psi_j u'_j, \quad (3.37)$$

for all  $j$ . Differentiating twice provides

$$H_{j_1 j_2} = v'' \psi_{j_1} u'_{j_1} \cdot \psi_{j_2} u'_{j_2} + v' \psi_{j_1 j_2} u'_{j_1} u'_{j_2} + \delta_{j_1 j_2} v' \psi_{j_1} u''_{j_1}, \quad (3.38)$$

where  $\delta_{j_1 j_2}$  is the Kronecker symbol and  $H_{j_1 j_2}$  abbreviates the second partial derivative of  $H$  with respect to  $x_{j_1}$ ,  $x_{j_2}$ . Expressing  $u'_j$  by the normal form constraint (3.37) and evaluating (3.38) for  $j_1 \neq j_2$  we obtain

$$H_{j_1 j_2} = \frac{1}{v'} \left( \frac{v''}{v'} + \frac{\psi_{j_1 j_2}}{\psi_{j_1} \psi_{j_2}} \right). \quad (3.39)$$

By summation over all  $n(n-1)$  pairs  $j_1 \neq j_2$ , the normal form constraint  $\sum H_{j_1 j_2} = 0$  of (3.23) implies

$$0 = n(n-1) \frac{v''}{v'} + \sum_{j_1 \neq j_2} q_{j_1 j_2}, \quad \text{i.e.} \quad (3.40)$$

$$\frac{v''}{v'} = -\bar{q}. \quad (3.41)$$

Here we have used  $H_{j_1 j_2} = 0$ , by (3.20), and we have abbreviated

$$\begin{aligned} q_{j_1 j_2} &:= \frac{\psi_{j_1 j_2}}{\psi_{j_1} \psi_{j_2}} \\ \bar{q} &:= \frac{1}{n(n-1)} \sum_{j_1 \neq j_2} q_{j_1 j_2}. \end{aligned} \quad (3.42)$$

To determine  $v'$  itself, we invoke the normal form condition (3.33) which, in view of (3.39), (3.41), implies

$$v' = \frac{v''}{v'} + q_{12} = q_{12} - \bar{q}. \quad (3.43)$$

We therefore obtain the *invariants*

$$I_{j_1 j_2}(\psi, u^0) := H_{j_1 j_2} = \frac{q_{j_1 j_2} - \bar{q}}{q_{12} - \bar{q}}, \quad (3.44)$$

with  $q_{j_1 j_2}$ ,  $\bar{q}$  defined in (3.42). Clearly  $\sum_{j_1 \neq j_2} I_{j_1 j_2} = 0$ ,  $I_{12} = 1$  are normalized, and  $I_{j_1 j_2} = I_{j_2 j_1}$ . This leaves precisely  $d(n, p) = \frac{1}{2}n(n-1) - 2$  independent invariants of order  $|k| \leq 2$  in  $\psi$ , as predicted in (3.34). Moreover, the equations

$$I_{j_1 j_2}(\psi, u^0) = I_{j_1 j_2}(h, x^0), \quad (3.45)$$

which hold by theorem 3.6, can be used to solve the identification problem of the black-white-black tree; see (3.31).

We emphasize that all derivatives of  $u_j$  and  $v$  had to drop out of the representations of the  $d(n, p)$  invariants, because these transformations are defined uniquely by  $\psi$  itself. Moreover the order of the required  $\psi$ -derivatives does not exceed the order of the invariants themselves.

Alternatively to the direct and elementary calculation above we could have retraced the explicit construction of the normalizing transformations  $u_j$  and  $v$ , as specified in the proof of lemma 3.1 and corollaries 3.2, 3.3. We briefly list some explicit results, which follow from such calculations in a straightforward manner.

For  $n = 3$  variables, table 3.1 indicates one free invariant of order  $p = 2$  and six additional invariants of order  $p = 3$ . Denoting the variables  $x_1, x_2, x_3$  by  $x, y, z$ , to avoid sub-subscripts, the normal form conditions (3.20), (3.23) for  $p = 2$  imply

$$H_{xy} + H_{yz} + H_{xz} = 0 \quad (3.46)$$

for the normal form  $H$ . In terms of  $\psi$ , and in the notation of (3.35), the invariant term  $H_{xy}$  is given by

$$H_{xy} = \frac{v'(0)}{3\psi_1\psi_2\psi_3} (2\psi_3\psi_{12} - \psi_1\psi_{23} - \psi_2\psi_{13}). \quad (3.47)$$

We recall that  $\psi$ -indices denote partial derivatives here and below:  $\psi_1 = \psi_{u_1}$ ,  $\psi_{12} = \psi_{u_1u_2}$ , etc.. The term  $H_{xz}$ , arising from the substitution  $y \leftrightarrow z$ , is obtained by correspondingly substituting  $2 \leftrightarrow 3$  in the  $\psi$ -derivatives of (3.47). Of course the relation (3.46) is now obvious from the explicit form (3.47). Picking  $k_0 = (1, 0, 1)$  in the normalizations (3.16), (3.22) of corollaries 3.2, 3.3 obviously eliminates  $v'(0)$ . We may then choose

$$I_{12} := \frac{H_{xy}}{H_{xz}} = \frac{2\psi_3\psi_{12} - \psi_1\psi_{23} - \psi_2\psi_{13}}{2\psi_2\psi_{13} - \psi_1\psi_{23} - \psi_3\psi_{12}} \quad (3.48)$$

as the only remaining invariant of order  $p = 2$ ; see table 3.1.

To determine the three functions  $u_1(x), u_2(y), u_3(z)$  in the argument of  $\psi$  from the invariant equalities (3.30) of theorem 3.6, we need at least two more invariants of order  $p = 3$ . These can easily be constructed from the following two explicit expressions. First we compute

$$\begin{aligned} H_{xyz} := \frac{v'(0)^2}{24\psi_1^3\psi_2^3\psi_3^3} \cdot \sum_{\text{perm}} & \left( 3\psi_1^2\psi_2^2\psi_3^2\psi_{123} - \right. \\ & - \psi_1\psi_2\psi_3(\psi_1^2\psi_{23}^2 + 2\psi_1\psi_2\psi_{13}\psi_{23}) - \\ & \left. - 3\psi_1^3\psi_2^2(\psi_3\psi_{233} - \psi_{23}\psi_{33}) \right). \end{aligned} \quad (3.49)$$

Here the summation runs over all permutations of the subscript indices  $\{1, 2, 3\}$ .

Subscripts again indicate partial derivatives of  $\psi$ . Similarly

$$\begin{aligned}
H_{xxy} = \frac{v'(0)^2}{144\psi_1^3\psi_2^3\psi_3^3} & \left( 9\psi_1^2\psi_2^2\psi_{33}(\psi_1\psi_{23} + \psi_2\psi_{13}) + \right. \\
& + 9\psi_1^2\psi_3^2\psi_{22}(\psi_1\psi_{23} + \psi_3\psi_{12}) + \\
& + 9\psi_2^2\psi_3^2\psi_{11}\psi_{13} - 63\psi_2^2\psi_3^3\psi_{11}\psi_{12} + \\
& + \psi_1\psi_2\psi_3(10\psi_1^2\psi_{23}^2 - 9\psi_1^2\psi_3\psi_{223} - \\
& \quad - 4\psi_1\psi_3\psi_{12}\psi_{23} - 14\psi_3^2\psi_{12} - \\
& \quad - 9\psi_1\psi_3^2\psi_{122} + 10\psi_2^2\psi_{13}^2 - \\
& \quad - 9\psi_1\psi_2^2\psi_{133} - 9\psi_2^2\psi_3\psi_{113} - \\
& \quad - 9\psi_1^2\psi_2\psi_{233} + 20\psi_1\psi_2\psi_{13}\psi_{23} - \\
& \quad - 4\psi_2\psi_3\psi_{12}\psi_{13} + 63\psi_2\psi_3^2\psi_{112}) - \\
& \left. - 18\psi_1^2\psi_2^2\psi_3^2\psi_{123} \right) \tag{3.50}
\end{aligned}$$

provides six additional terms, by the associated permutations of  $\{x, y, z\}$ ,  $\{1, 2, 3\}$ , respectively. In all, and after the elimination of  $v'(0)$  at second order, say by  $H_{xz} = 1$ , and in view of the normal form conditions (3.20), (3.23), we obtain six additional independent invariants from (3.49), (3.50), at order  $p = 3$ . The total  $d(3, 3) = 7$  of table 3.1 agrees with this derivation. As a perhaps more manageable invariant we mention the following antisymmetric form of (3.50):

$$H_{xxy} - H_{xyy} = \frac{v'(0)^2}{2\psi_1^3\psi_2^3} (\psi_1^2\psi_{12}\psi_{22} + \psi_1\psi_2^2\psi_{112} - \{12\}) \tag{3.51}$$

together with its two other substitution associates. Here the term  $\{12\}$  repeats all other terms inside the parenthesis with derivative indices 1, 2 of  $\psi$  interchanged.

As our final example, we consider  $n = 2$  variables. Starting at order  $p = 3$ , as suggested by table 3.1, we obtain

$$H_{xxy} = \frac{v'(0)^2}{4\psi_1^2\psi_2^2} (\psi_1^2(\psi_{12}\psi_{22} - \psi_2\psi_{122}) - \{12\}). \tag{3.52}$$

Again,  $\{12\}$  repeats the preceding term with 1, 2 interchanged. The obvious anti-

symmetry follows from (3.20), (3.23). At order  $p = 4$ , we obtain

$$\begin{aligned}
H_{xxyy} = \frac{v'(0)^3}{14\psi_1^3\psi_2^3} & \left( -3\psi_1^4\psi_{12}\psi_{22}^2 + \right. \\
& + \psi_1^3\psi_2(3\psi_1\psi_{22}\psi_{122} + \psi_1\psi_{12}\psi_{222} + 2\psi_{12}^2\psi_{22}) + \\
& + \psi_1^2\psi_2^2(\psi_{11}\psi_{22} + \psi_{1222} - 2\psi_1\psi_{22}\psi_{112} - 2\psi_1\psi_{12}\psi_{122}) + \\
& \left. + \{12\} \right)
\end{aligned} \tag{3.53}$$

in a similar notation. For example, we can use (3.53) and the normalization  $H_{xxyy} = 1$  to eliminate  $v'(0)$ . The second independent term up to order  $p = 4$  can then be chosen to be the antisymmetrized expression

$$\begin{aligned}
H_{xxyy} - H_{xyyy} = \\
\frac{v'(0)^3}{6\psi_1^5\psi_2^5} & \left( \psi_1^4(-3\psi_{12}\psi_{22}^2 + \psi_2(\psi_{12}\psi_{222} + 3\psi_{22}\psi_{122}) - \psi_2^2\psi_{1222}) - \{12\} \right).
\end{aligned} \tag{3.54}$$

For any number  $n \geq 2$  of inputs to the black-white-black tree of Figure 2.1, the above explicit expressions (3.44), (3.48)–(3.54) thus provide at least  $n$  invariant determining equations in terms of the white box  $\psi$  and the given data  $h$  for the  $n$  unknown black-box functions,  $u_1, \dots, u_n$ . Since the  $n$  determining equations can be regarded as independent, in general, this provides a method for the local identification of the black-box functions. For generically unique identification of the black boxes  $u_j$  we will require a slightly larger total number of  $2n + 1$  determining equations, in the next section.

Concerning the extrapolability of the above identification of the elementary bwb-tree of Figure 2.1 we recall our definition (1.17) of the spanned data base  $\overline{D}$ . Consider any open arc of the one-dimensional data base curve  $D$ , replacing  $D$  itself. Then the input sets  $D_j$  to the black boxes  $u_j, u_j = 1, \dots, n$ , are open intervals. By monotonicity (1.10), the input set to the final black box  $v$  is also an open interval spanned by the values at the end points of  $D$ . Therefore

$$\overline{D} = D_1 \times \dots \times D_n \tag{3.55}$$

is the  $n$ -dimensional quadrangle spanned by the intervals  $D_j, j = 1, \dots, n$ .

## 4 Genericity

In this section we proceed towards the proof of our main theorem 1.1 for the elementary black-white-black tree of section 3. Our proof is based on systems of equations

$$I_k(\psi, u^0) = I_k(h, x^0) = H_k \quad (4.1)$$

for normal form invariants  $H_K$ ; see theorem 3.6. We evaluate (4.1) for multi-indices  $k$  in a suitable index set  $K \subset \{0, 1, \dots, \nu\}^n$  of cardinality  $|K| = 2n + 1$ . We claim global uniqueness of solutions

$$u^0 = u^0(x^0) \quad (4.2)$$

of system (4.1), for given global input-output data  $h$ , and for generic white-box functions  $\psi \in X \subset C^\nu(\mathbb{R}^n, \mathbb{R})$ . For a topology on  $C^\nu$  we choose the weak Whitney topology of  $C^\nu$ -uniform convergence on compact subsets. The subset  $X$  is defined by the condition of nonzero partial derivatives  $\psi_{u_j}$ ; see assumption (1.10). As usual, genericity is understood in the Baire sense: uniqueness of  $u^0(x^0)$  will be shown to hold for  $\psi$  in a residual subset of  $X \subset C^\nu$ , viz. at least for  $\psi$  in a countable intersection of open dense sets. See [14] and [1] for a background.

To prove our claim, we show that the map

$$\begin{aligned} I &:= (I_k)_{k \in K} : \mathbb{R}^n \rightarrow \mathbb{R}^{2n+1}, \\ u^0 &\mapsto (I_k(\psi, u^0))_{k \in K} \end{aligned} \quad (4.3)$$

is injective, for generic white-box functions  $\psi \in X$ . As usual, for example in the proof of the Whitney embedding theorem, we consider the associated map

$$\begin{aligned} \mathcal{I} &: \tilde{X} \times (\mathbb{R}^{2n} \setminus \Delta) \rightarrow \mathbb{R}^{2n+1}, \\ (\psi, u^1, u^2) &\mapsto I(\psi, u^1) - I(\psi, u^2). \end{aligned} \quad (4.4)$$

Here  $\tilde{X} \subseteq X$  is an open dense subset, and  $\Delta$  denotes the diagonal  $u^1 = u^2$  of pairs  $(u^1, u^2) \in \mathbb{R}^{2n}$ . Note how non-injectivity of  $I$  corresponds to zeroes of  $\mathcal{I}$ . Injectivity of  $I$  is proved, if  $\mathcal{I}(\psi, \cdot, \cdot)$  does not attain the value zero on  $\mathbb{R}^{2n} \setminus \Delta$ .

In lemma 4.1 below we prove that zero is a regular value of the map  $\mathcal{I}$ , for suitable choices of the index set  $K$ . By Thom's transversality theorem, alias Sard's theorem, this implies that zero is a regular value of the map

$$\mathcal{I}(\psi, \cdot, \cdot) : \mathbb{R}^{2n} \setminus \Delta \rightarrow \mathbb{R}^{2n+1} \quad (4.5)$$

for generic white-box functions  $\psi \in \tilde{X}$ . But the partial derivative of  $\mathcal{I}(\psi, \cdot, \cdot)$  cannot be surjective, for any fixed  $\psi$ , by dimensional reasons. Therefore zero, being a regular value of  $\mathcal{I}(\psi, \cdot, \cdot)$ , cannot possibly be a *value* of  $\mathcal{I}(\psi, \cdot, \cdot)$  — and injectivity of  $I$  is proved.

A minor technical point in the above argument is the requirement of a Banach space for the white-box function  $\psi$ . For example, we may restrict  $\psi$  to large balls  $u_1^2 + \dots + u_n^2 \leq R^2$ , where its positive partial derivatives  $\psi_{u_j}$  are still uniformly bounded away from zero, in addition to all derivatives of order up to  $\nu$  being uniformly bounded. The corresponding set is then open in the strong topology  $BC^\nu$  of  $C^\nu$ -uniform convergence in sup-norm. To prove the residual property in the weak Whitney topology, we can then let  $R$  tend to  $\infty$ . For an alternative, abstract version of Thom's result, which applies directly to the present situation, see [7]. Below we keep this small technicality in mind, when writing  $\psi \in X \subseteq C^\nu$  and referring to open dense subsets.

**Lemma 4.1** *Choose  $\nu \geq \nu(n)$  large enough. Then there exist index sets  $K \subseteq \{0, 1, \dots, \nu\}^n$  of cardinality  $2n + 1$  and an open dense subset  $\tilde{X} \subseteq X \subseteq C^\nu$ , such that the map  $\mathcal{I}$  defined in (4.4) possesses a surjective partial derivative with respect to  $\psi$ , at every point  $(\psi, u^1, u^2) \in \tilde{X} \times (\mathbb{R}^{2n} \setminus \Delta)$ .*

**Proof:** We will address differentiability issues first. We then reduce surjectivity of the partial derivative  $\mathcal{I}_\psi$  to surjectivity of  $I_\psi$ . By the group action (3.1) of the transformations  $v, u_j$ , we reduce surjectivity of  $I_\psi$  to the special case where  $\psi$  is already in normal form:  $\psi = H$ . By linearity of the space of normal forms in the jet space, surjectivity at  $\psi = H$  will be trivial, for proper choices of the index set  $K$  and of the open dense subset  $\tilde{X} \subset X$  of white-box functions  $\psi$ .

Concerning differentiability, we first observe that the unique explicit  $\nu$ -jets of the transformations  $u, v$  to normal form  $H$  depend  $C^1$  on  $\psi \in C^\nu$  for  $\nu \geq 2$ ; see (3.7), (3.9)–(3.12) and the normalization (3.18) of (3.15) to (3.16). For the step (3.11) of linearization of  $\varphi$  by  $\varphi(v(t)) = v(nt)$ , this follows from the explicit convergent representation of the inverse  $v^{-1}(t) = t + \tilde{v}(t)$  by the convergent sum

$$\tilde{v}(t_0) = - \sum_{j=1}^{\infty} n^{j-1} \tilde{\varphi}(t_{-j}). \quad (4.6)$$

Here we have also decomposed  $\varphi(t) = nt + \tilde{\varphi}(t)$ , and we denote the backwards iteration of  $\varphi$  by  $t_{-j} = \varphi^{-j}(t_0)$ . Transforming in the restricted class  $v'(\psi(u^0)) = 1$ , we also see that the nondegeneracy condition  $H_{k_0} \neq 0$  defines an open dense subset  $\tilde{X}$  of  $X$ . In fact, the normal form  $H \in C^\nu$  of corollary 3.3, with the normalization  $H_{k_0} = 1$ , then depends  $C^1$  on  $\psi \in \tilde{X}$ . This proves  $\mathcal{I} \in C^1$ .

Next suppose  $I = I(\psi, u)$  of (4.3) possesses a surjective partial derivative  $I_\psi$  at any  $\psi = \psi^0 \in \tilde{X}$ ,  $u = u^0 \in \mathbb{R}^n$ . Then

$$\mathcal{I}_\psi(\psi, u^1, u^2) = I_\psi(\psi, u^1) - I_\psi(\psi, u^2) \quad (4.7)$$

is also surjective at any  $\psi = \psi^0 \in \tilde{X}$  and for any off-diagonal pair  $(u^1, u^2) \in \mathbb{R}^{2n} \setminus \Delta$ . Indeed, the assumed surjectivity of  $\tilde{\psi} \mapsto I_\psi(\psi, u^\iota)\tilde{\psi}$ ,  $\iota = 1, 2$  is achieved by the  $\nu$ -jets of  $\tilde{\psi}$  at the two different points  $u = u^1, u^2$ , separately, and thus combines to provide surjectivity of  $\mathcal{I}_\psi$ .

To prove surjectivity of  $I_\psi(\psi, u)$ , at any  $\psi \in \tilde{X}$ ,  $u \in \mathbb{R}^n$ , we assume  $\psi$  to be in normal form,  $\psi = H$ . This restriction does not constitute any loss of generality. Indeed, small neighborhoods of the  $\nu$ -jet  $J^\nu\psi(u)$  are  $C^1$ -parametrized by the  $\nu$ -jet  $J^\nu H(0)$  of the normal form  $H$ , and by the  $C^\nu$ -transformations  $u_j, v$ . In normal form, however,  $I(\psi, u)$  is itself linear with respect to  $\psi$  within the normal form space. It is therefore automatically surjective, if the index set  $K$  of cardinality  $2n + 1$  avoids the restrictions imposed by the normal form constraints (3.20)–(3.23) of corollary 3.3. This completes the proof of the lemma.  $\times$

## 5 Proof of theorem 1.1

In this section we summarize the proof of theorem 1.1, as prepared in sections 2-4. We recall the reduction of the identification problem to black-white-black trees, of section 2, as well as the identification in section 3 for generic intermediate white boxes  $\psi = u_j$ ,  $j \in W$ , of section 4. Together, this will prove theorem 1.1.

In sections 2–3 we have shown how to identify any black boxes  $u_j$  in the first layer,  $j \in \ell_1 \cap B$ , turning them white. This identification was unique only up to calibration, in case of subsequent black boxes in layer  $\ell_2$ . Identification was based on the Frobenius



type lemma 2.1. This lemma, in turn, used the monotonicity assumption (1.10) and, via the connectedness lemma 2.2, the minimal transversality assumption (1.15) as well as assumption (1.16) on connectedness of level sets in the minimally transverse data base  $D$ . In the case of a subsequent white box  $u_{s(j)}$ , unique identification of the black box  $u_j$  was completed by the normal form systems (3.30) of section 3, with the genericity and injectivity considerations of section 4, (4.1)–(4.5). This step uses the genericity assumption (1.11) on white boxes  $\psi = u_j$ ,  $j \in W$ ; see lemma 4.1.

From table 3.1 we conclude that the maximal order  $\nu$  of derivatives necessary to capture  $2n + 1$  independent invariants  $H_k$  is  $\nu = 5$ , which arises at white boxes  $\psi$  with only  $n = 2$  inputs.

These arguments complete the proof of theorem 1.1 for the first layer,  $j \in \ell_1$ . To repeat the same steps for the remaining layers, we consider the outputs  $x_j$ ,  $j \in \ell_1$ , as new input variables. We can then repeat the same identification process for input vectors  $x^{j'} = (x_j)_{j \in p(j')} \in D_{j'}$  to black boxes  $j'$  in layer  $\ell_2$  and so on, recursively, through the scalar hybrid tree.

We repeat here that each identification step of the successive layers  $\ell_j$  is based on the global input-output relation  $h$  which results, when all black boxes of previous layers have been identified and thus have turned white. We can therefore consider all boxes in the layer  $\ell_j$  as first layer boxes, for the induction step. Note how transversality of the data base  $D$  and connectedness of the sets  $D_{j,0}^c$  is inherited during this replacement. The inputs for black boxes in layer  $\ell_j$  are therefore sufficient for identification of the SHM on the full spanned set  $\bar{D}$  of inputs, as defined in (1.17). This proves theorem 1.1.

## 6 Industrial application

As an example for an industrial application, we describe a SHM for a continuous polymerization plant. Polymer materials consist of long, highly entangled molecular chains which can crystallize by folding of the chains. The plant model has been used to optimize and control the “melt index” (MI), an important quality property for polymers depending on the chain length distribution of the product [10]. The

material produced in the process under consideration have special material structures, as are required by technically demanding applications. The entire process has a feed-forward structure. Monomers and co-monomers polymerize in a tube reactor, started by a catalyst. The chain length distribution in the reactor is controlled by an additional species which stops the polymerization at the end of a polymer chain and starts the polymerization reaction for another chain. Because the reaction in the reactor under consideration is accompanied by crystallization processes with widely unknown impacts on the reaction rates; all attempts of rigorous modeling for this reaction process had failed. The resulting raw polymer which leaves the reactor is subsequently separated from rests of educts, stabilized, and granulated in separate process steps. These steps do not influence the chain length distribution significantly, and have been modeled by white-box mixing models. Recycles of energy and materials do not influence the melt index MI either. In industrial practice the continuous plants are always operated near the expected optimal setpoints for the process parameters. The only changes in the process parameter space are performed, if the throughput of the plant or the quality (MI) of the product have to be changed. In both cases the control of the reaction is performed by correlated changes in the chain length controller concentration and in the catalyst concentration. The resulting high correlation in the process parameters had caused the training of pure black-box models to fail. Therefore optimization and model based control of MI could only be performed using a SHM for the reactor.

Using the chemical know-how about the reactive part, a SHM with the following structure has been developed. The melt index MI depends on the chain length distribution of the polymers. For amorphous polymers the dependence of the chain length distribution on the average chain length  $c_L$  is well established. In our case, however, the respective relation had not been established with the necessary reliability. Therefore a black-box model had to be used for the dependence  $w$  of the melt index MI on the average chain length  $c_L$ :

$$\text{MI} = w(c_L) \tag{6.1}$$

The average chain length  $c_L$  itself is controlled by the yield  $Y$  of the polymerization reaction with respect to the monomer feed and the number  $N_c$  of chains produced

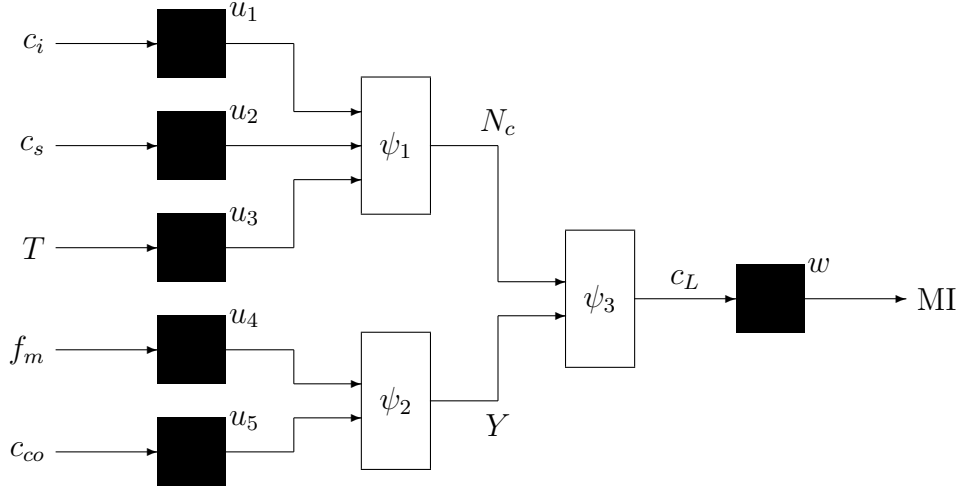


Figure 6.1: SHM for the polymerization mechanism. The vector  $c = (c_i, c_s, c_{co}) \in \mathbb{R}^3$  describes the concentrations of initiator, chain length controller, and comonomer species;  $f_m$  denotes the monomer flow and  $T$  the feed flow temperature.

per mole of monomers:

$$c_L = \psi_3(Y, N_c) \quad (6.2)$$

Therefore, the effect of the process input variables can be separated into two (abstract) subprocesses for the yield  $Y$  and the number  $N_c$  of chains per mole. Both  $Y$  and  $N_c$  are input variables for the white-box subprocess  $\psi_3$  of (6.2) in which the average chain length  $c_L$  is calculated. This separation of the black-box subprocesses strongly depends on the qualitative understanding of the chemical polymerization mechanism.

Using additional chemical information about the structure of the reaction network we arrive at a further reduction of the complexity of the SHM for the reactor subprocesses as depicted in Figure 6.1.

The SHM mapping the inputs  $c = (c_i, c_s, c_{co})$ ,  $T$ ,  $f_m$  to the output MI for the reactive part of the process therefore takes the form:

$$\begin{aligned} N_c &= \psi_1(u_1(c_i), u_2(c_s), u_3(T)) \\ Y &= \psi_2(u_4(f_m), u_5(c_{co})) \\ \text{MI} &= w(\psi_3(Y, N_c)) \end{aligned} \quad (6.3)$$

with unknown black-box functions  $u_1, \dots, u_5$  and  $w$  depending on only one input

variable each. The processes  $\psi_1$  and  $\psi_2$  have been modeled by known white boxes. Moreover, the input variable  $c_L$  to the melt-index black box  $\text{MI} = w(c_L)$  has been replaced by the white-box output  $c_L = \psi_3(N_c, Y)$ .

Alternatively, the three white boxes  $\psi_1, \psi_2, \psi_3$  could have been combined into a single white box  $\psi$  with five inputs. Then our SHM takes the precise black-white-black form studied in section 3.

Applying theorem 1.1 above to this model structure, instead of the alternative trivial black-box approach with an unknown function of five input variables:

$$\text{MI} = u(c, T, f_m) : \mathbb{R}^5 \rightarrow \mathbb{R}, \quad (6.4)$$

the dimension of the required data base has been reduced from five to one. This dramatically reduces the data demand. Even more importantly, the SHM model can be used for extrapolations even when it has been trained on highly correlated data sets, only. The model has been used to fit the melt index of the polymerization plant with respect to the input variables.

Alternatively, a pure black-box model, realized by a neural network, has been trained on the same data. For the underlying real process, the neural network alone requires about 2000 data sets to predict the melt index properly, whereas the SHM can be used for melt-index prediction after training on only 200 data sets. Moreover, the prediction quality of the SHM is about five percent higher than that of the neural network, although the training set has been 10-fold smaller. Moreover, due to the low dimension of the data base, the SHM allows the analysis of the impact of monomer impurities on the melt index from plant data without any additional experiments.

So, how have all the assumptions of theorem 1.1 been verified in the application to the polymerization model? The data base, first of all, cannot really be “chosen”. It consists of only such input data, in the industrial context, for which the overall output variable MI has in fact been measured. We are restricted to use this data, only, as other data are not available. By interpolation, rather than extrapolation, we then have extracted the nontrivial second order invariants, as required by section 2 for  $n \geq 5$  inputs and  $p = 2$ , along a one-dimensional curve in input data space; see (3.37) and (3.44), (3.45). By (3.34) there are eight such invariants. Monotonicity (1.10) does hold for the specific white-box models.

Already in the introduction we have pointed out the crucial relevance of the curse of dimension and of the requirement of extrapolability for practical reactor modeling. Indeed and quite generally, data is generated only under correlated changes of operating conditions. This practical constraint leads to data distributions which are clustering around a low dimensional subset in the space of process parameters. This low dimensional subset is often believed to be optimally designed to run the process with respect to the available educts and the desired product specifications. Because of this (pseudo-)optimality the dimension of the data manifold is usually low and the adjustment of inputs follows clear and simple rules for plant operation. These practical restrictions, in turn, lead to highly monotone co-variations in the data bases available for the overall process.

With only a one-dimensional data base required in the present example, our approach certainly avoids the curse of dimension. Moreover, extrapolability comes for free with the identification of our structured hybrid model.

That genericity (1.11) cannot be checked, a priori, is a price to be paid for such advantages. In fact, we just solve the resulting determining equations for the small black boxes  $u_1, \dots, u_5$ . Global uniqueness is asserted for most white-box functions  $\psi$ . This is an indication that identification should *mostly* work. Even in absence of trivial calibrations, though, global uniqueness of the black box solutions cannot be asserted for *all* white-box functions  $\psi$ , in general. See our simple introductory example (1.12), (1.13).

The genericity problem of our application has appeared in the literature before. It is reminiscent of the justly celebrated Takens embedding result [19]: for generic diffeomorphism  $\phi$  of a compact manifold  $\mathcal{M}$ , and for generic maps  $P : \mathcal{M} \rightarrow \mathbb{R}$ , the sequence  $(P\phi^n(x))_{n \in \mathbb{N}}$  of scalar iterates allows a reconstruction of the dynamics  $\phi$  on  $\mathcal{M}$ . More precisely, the delay embedding map  $\mathcal{J} : \mathcal{M} \rightarrow \mathbb{R}^m$  given by  $\mathcal{J}(x) := (P\phi^m(x), \dots, Px)$  becomes an embedding for  $m \geq 2 \dim \mathcal{M} + 1$ . Genericity of  $P, \phi$  cannot be checked explicitly, from knowledge of  $P$  and the resulting dynamics  $\mathcal{J}(\phi^n(x))_{n \in \mathbb{N}}$  alone. Still, the genericity assumption cannot be discarded, as the simple example  $\phi = \text{id}$  shows:  $\mathcal{J}(x)$  is then constrained to lie in the one-dimensional diagonal of  $\mathbb{R}^m$ , for any  $x \in \mathcal{M}$ . An analogously mild, but necessary, caveat applies to our genericity assumption in (1.11). In this sense, on the other hand, our generic-

ity assumption strongly supports educated optimism as to the universal applicability of our result.

In a final step, after proper identification of all black boxes  $u_j$ , the foliation of the input space  $x^0 \in \mathbb{R}^5$  by output values MI is known. The minimal transversality assumption (1.15) as well as the connectedness assumption (1.16) can then be checked, a posteriori, for the specific system.

Should global transversality fail, we can still decompose the input space into large regions where the components of the data base satisfy the transversality conditions, respectively and locally. In each of these transversality regions, separately, our theorem applies. If more than one such transversality region occurs, overlap of the respectively resulting extrapolation regions may result. Any such overlap then offers an additional opportunity for a – highly welcome – global check of the model for reliability and data consistency.

Together, the above steps complete the unique identification of the SHM on the extended, fully five-dimensional data base  $\bar{D} \subset \mathbb{R}^5$  from measured data clustered along the original curve  $D \subset \mathbb{R}^5$ .

For further and more detailed examples we again refer to the technical report [11].

## 7 Summary and discussion

We have considered structured hybrid models (SHM), which are trees  $j \in \{1, \dots, N\}$  with successor  $s(j)$  and predecessors  $p_j = s^{-1}(j)$  of the vertex, or box,  $j$ .

Assuming monotonicity of the scalar box functions  $u_j$ , genericity for white boxes  $j \in W$ , and inputs on a  $d$ -dimensional, minimally transverse data base  $D$  of dimension  $d = \max_{j \in B} |p_j|$ , we have uniquely identified all black boxes  $u_j$ ,  $j \in B$ , up to trivial calibrations.

We have not addressed the delicate question of existence of solutions  $u_j$ ,  $j \in B$ , for given input-output data  $h$ . In fact the results by Vitushkin, mentioned in the introduction, indicate that typical data  $h$  cannot be realized by *any* black-box hybrid model of tree type, if  $h$  possesses more inputs in total than any single one of the

black boxes  $u_j$ , and if an equal order of differentiability is required for, both, data  $h$  and solutions  $u_j$ . It may therefore be a quite delicate task to provide a useful description of the range of all input-output data functions  $h$  which are generated by a particular hybrid model. Certainly the abundance of transformation group invariants  $I_k$  given by normal form coefficients  $H_k$ , for the black-white-black tree of section 3, and the resulting grossly overdetermined systems

$$I_k(\psi, u^0) = I_k(h, x^0) \quad (7.1)$$

already provide many consistency checks to be matched by the data  $h$ .

A related problem is the stability of numerical algorithms, which actually determine the unique black-box solutions  $u_j$ ,  $j \in B$ , in a given scalar hybrid tree. As it stands, our local invariant systems (7.1), for example, at best provide solutions  $u^0 = u^0(x^0; \psi, h)$  which depend differentiably on  $\psi, h \in C^\nu$ ,  $\nu \geq 4$ . Moreover, different choices of index sets  $k \in K$ , say of cardinality  $n$ , may lead to different solutions  $u_K^0(x^0; \psi, h)$  with different numerical condition numbers for the solutions of (7.1). Dependence of the solution on  $K$  then indicates inconsistency of the data.

The high differentiability requirements on  $\psi$ ,  $h$  present a formidable obstacle to any naive application of our result in the setting of (even very large sets of) point measurements of input-output data  $h$ . It is conceivable, however, that nonlocal identification procedures, for example based on differential equations for the black-box functions  $u_j$ ,  $j \in B$ , help alleviate this problem.

The monotonicity requirement (1.10) is essential to our uniqueness result. For example, we have numerically investigated a simple two layer hybrid model

$$\psi(u_1(x_1), u_2(x_2)) = h(x_1, x_2) \quad (7.2)$$

with a nonmonotone, polynomial white box  $\psi$  of cubic order; see [11]. We have encountered up to 11 solutions  $u_j^0$  of the associated invariant system

$$\begin{aligned} \psi(u_1^0, u_2^0) &= h_1^0 \\ \frac{\psi_{u_1 u_2}}{\psi_{u_1} \psi_{u_2}}(u_1^0, u_2^0) &= h_2^0. \end{aligned} \quad (7.3)$$

A serious structural limitation of our result is imposed by the scalar nature  $u_j \in \mathbb{R}$  of the outputs. For general vector couplings  $u_j \in \mathbb{R}^{n_j}$ , nontrivial scalings may arise. In

the crucial vector black-white-black example, analogously to section 3, the normal form approach looks promising, at least as long as the black boxes  $v, u_j$  remain diffeomorphisms.

Finally we mention the case of hybrid models with feedbacks, alias cycles in the model graph  $\Gamma$ . This case, so particularly relevant both in chemical industry and for biological systems, has not even been touched from the view point of identification of hybrid models which we have presented here.

Even so, our preliminary results already open wide areas of applications which have been inaccessible before, in an industrial framework. In contrast to black-box or unstructured hybrid models, using only one big black-box submodel and lumping all unknown subprocesses, the graph structure of the SHM exhibits three significant advantages over black-box and rigorous models alike:

- (i) The dimension of data bases can be reduced. This allows the identification of the model from highly correlated data as well as a significant reduction of the data demand to fit the model, without loss of accuracy. This in turn reduces the latency time for model-based on-line applications significantly. For an on-line polymerization plant, the data demand for the fit of the model has been reduced by 90%.
- (ii) While the dimension of the required data base is reduced, the SHM remains extrapolable. In contrast to black-box models with nearly total lack of extrapolability, SHMs can therefore be applied for control and optimization.
- (iii) The extrapolability of the SHM allows its use in process analysis, on-line investigating for example the effect of impurities or other disturbances on the product quality. In case of any process disturbance, the operators will try to keep the product quality constant through changing appropriate control parameters. Therefore the values of the process disturbances and control parameters are closely related. A retrospective analysis of the process dependence on the disturbances by only black-box models is impossible. Appropriate SHMs, however, do allow extrapolation from correlated data, and can provide a rigorous mathematical foundation for a rational analysis of the relation between product quality and process disturbances.



## References

- [1] R. Abraham and J. Robbin. *Transversal Mappings and Flows*. Benjamin Inc., Amsterdam, 1967.
- [2] V. I. Arnol'd. On functions of three variables. *AMS Transl.*, **28**:51-54 (1963).
- [3] V. I. Arnol'd. *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer-Verlag, New York, 1983.
- [4] V. I. Arnol'd (ed.). *Singularity Theory I*. Enc. Math. Sc. **6**, Springer-Verlag, Berlin, 1993.
- [5] G. R. Belitskii. Functional equations and conjugacy of local diffeomorphisms of a finite smoothness class. *Funct. Analysis Appl.* **7**:268–277 (1973).
- [6] H. A. B. Te Braake, H. J. L. van Can and H. B. Verbruggen. Semimechanistic modeling of chemical processes with neural networks. *Eng. Appl. Art. Intell.*, **11**:507-515 (1998).
- [7] J. Damon. Generic properties of solutions to partial differential equations. *Arch. Rat. Mech. Analysis*, **140**:353-403 (1997).
- [8] M. Dors, R. Simutis and A. Lübbert. Hybrid process modeling for advanced process state estimation, prediction and control exemplified in a production-scale mammalian cell culture. *ACS Symposium Series*, **613**:144-154 (1995).
- [9] J.-P. Dufour. Sur la stabilité des diagrammes d'applications différentiables. *Ann. Sci. Éc. Norm. Supér., IV. Sér.*, **10**:153-174 (1977).
- [10] *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley, 1998.
- [11] B. Fiedler, S. Liebscher et al. *Strukturierte Hybridmodelle*. Abschlussbericht BMBF 2004, Freie Universität Berlin (German).
- [12] B. L. Friedman. An improvement in the smoothness of functions in A. N. Kolmogorov's theorem on superposition. *Soviet Math. Dokl.*, **8**:1550-1553 (1967).
- [13] D. Hilbert. Mathematische Probleme. *Archiv der Mathematik und Physik*, **1**:44-63 (1900).

- [14] M. W. Hirsch. *Differential Topology*. Springer-Verlag, New York, 1976.
- [15] A. N. Kolmogorov. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition. *Trans., II. Ser., AMS*, **28**:55-59 (1963).
- [16] D. C. Psichiogos and L. H. Ungar. A hybrid neural network - first principle approach to process modeling. *AIChE Journal*, **38**:1499-1511, (1992).
- [17] A. Schuppert. New approaches to data-oriented reaction modeling. In *Proceedings of the 3rd Workshop on Modeling of Chemical Reaction Systems*. Springer-Verlag, Heidelberg, 1996.
- [18] A. Schuppert. Extrapolability of structured hybrid models: A key to optimization of complex processes. In B. Fiedler et al., editor, *International Conference on Differential Equations*, pages 1135-1151. World Scientific Publ., Singapore, 2000.
- [19] Floris Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence*, D. Rand and L.-S. Young (eds.), Lect. Notes Math. **898**:366-381. Springer-Verlag, Berlin, 1981.
- [20] M. L. Thompson and M. A. Kramer. Modeling chemical processes using prior knowledge and neural networks. *AIChE Journal*, **40**:1328-1340 (1994).
- [21] A. Vanderbauwhede. Centre manifolds, normal forms and elementary bifurcations. In U. Kirchgraber and H. O. Walther, editors, *emphDynamics reported* **2**:89-169. Teubner & Wiley, Stuttgart, 1989.
- [22] A. G. Vitushkin. On Hilbert's thirteenth problem. *Dokl. Akad. Nauk SSSR*, **95**:701-704 (1954).
- [23] A. G. Vitushkin. On representation of functions by means of superpositions and related topics. *L'Enseignement Mathématique Ser. II.*, **23**:255-320 (1977).