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

May 2003

1 Introduction

By a *structured hybrid model* we mean a finite oriented Graph $\Gamma = \{1, \ldots, N\}$ equipped with certain functions u_j, associated to vertices $j \in \Gamma$, and with *global input-output data* h. The oriented edges of the graph Γ indicate composition of the vertex functions u_j, thus giving rise to h, as we will make precise in (1.8) below. Edges pointing towards vertex j are called *inputs* of u_j, whereas edges pointing away from j are called *outputs*. 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. In the present paper, we show how such information can be sufficient to identify all black boxes. Moreover, we indicate how all black boxes can be identified, successively and constructively.

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 precesses u_j of the black boxes $j \in B$.

Structured hybrid models arise frequently in problems of 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 for reactive subprocess. Therefore advanced modeling technologies 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 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 like neural networks, on the other hand, allow the quantitative description of the *input-output* (i/o) *relations* of the overall process, which is often sufficient for industrial use. 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.

(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 technologies which combine both approaches [TK94, PU92, BvCV98, DSL95]. 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. 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 describing the i/o-relation of those subprocesses which are well understood

(b) black-box submodels for those subprocesses for which no rigorous models are available

(c) a model flowsheet 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 only the simple reduction of complexity. The explicit integration of the process flowsheet structure 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-kinetic modeling [Sch96] and will be demonstrated in the present paper with an example of a continuous polymerization process; see section 6.

Early mathematical fame has been associated with a variant of our hybrid modeling problem. In the XIII. of his celebrated problems, Hilbert has asked to express the roots of higher order polynomials by compositions of scalar functions with at most two scalar input variables; [Hil00]. In other words, the graph Γ 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 [Arn63] and, in definitive form, by [Kol63]. 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

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

Here $\psi_{i,j} \in C^0(\mathbb{R}, \mathbb{R})$ are universally prescribed given C^0-functions. 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 [Fri67]. For related results in the category of functions $h \in C^\nu(\mathbb{R}^n, \mathbb{R})$, $u_j \in C^{\nu'}(\mathbb{R}^{n'}, \mathbb{R})$, see [Vit54], [Vit77]. One of these results relies on the *Vitushkin complexity* ν/n. In fact, most data h cannot be represented by any finite composition of functions u_j of higher Vitushkin complexity

$$\nu'/n' > \nu/n. \tag{1.2}$$

We repeat that the structure graph Γ, in these results, is not prescribed a priori, 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 Γ 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 to determine the black-box functions u_j from minimal information on h.

In the present paper, we restrict our attention to structure graphs Γ which are trees.

In other words, Γ is connected via the edges, and does not possess (unoriented) cycles. More specifically, there exists a map

$$\begin{array}{rrcl} s: & \Gamma \setminus \{N\} & \to & \Gamma, \\ & j & \mapsto & s(j) \end{array} \tag{1.3}$$

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 Γ into layers ℓ_k, $k = 0, \ldots, m$, defined by

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

The layer ℓ_0 will indicate the *inputs* to the structured hybrid model. Note that $\ell_m = \{N\}$. We also use the notation $\Gamma' := \Gamma \setminus \ell_0 = \mathrm{range}\,(s)$ 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 may be any positive integer, depending on j.

To each black or white box vertex $j \in \Gamma'$ we associate a scalar, ν times differentiable function $u_j \in C^\nu(\mathrm{I\!R}^{|p_j|}, \mathrm{I\!R})$ 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 the input to the hybrid model. The *input vector* is abbreviated by

$$x^0 := (x_j)_{j \in \ell_0}. \tag{1.6}$$

With this notation the composition of outputs and inputs along the tree structure Γ of the hybrid model takes the recursive form

$$x_j = u_j((x_i)_{i \in p_j}), \tag{1.7}$$

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 input vector $x^0 = (x)_{i \in \ell_0}$ the relation (1.7) indeed defines all outputs, recursively for j in increasing layers ℓ_k, $k = 1, \ldots, m$ of the tree Γ until finally

$$x_N = h(x^0). \tag{1.8}$$

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 functional, depending on the graph Γ, of the black and white-box functions u_j:

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

Our task is therefore to identify $(u_j)_{j\in B}$ from (1.9), under minimal information on h.

In particular the "curse of dimension" constrains h to be evaluated on 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 an embedded d-dimensional C^ν submanifold of $\mathbb{R}^{|\ell_0|}$. Moreover we require the ν-jet of the Taylor coefficients $D^i h(x^0)$, $|i| \leq \nu$, to be given on $x^0 \in D$.

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

$$u_{j,x_i} \neq 0 \tag{1.10}$$

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 $(x_i)_{i\in p_j} \in \mathbb{R}^{|p_j|}$. The nonmonotone case in fact poses serious obstacles to unique identification of black boxes which we plan to discuss elsewhere.

The data base D is assumed to be *in general position* and of dimension

$$d := \dim D = \max_{j\in B} |p_j|. \tag{1.11}$$

The assumption of "general position" is a *genericity assumption.* If we think of D as a graph over any d components of the input vector x^0, then the assumption holds for a countable intersection of open-dense sets in the C^ν Whitney topology; see [Hir76]. Here and below we understand genericity statements in this sense. For an explicit version of this condition on D see section 4 below.

Our final, third assumption concerns the global input-output function on the input-data $x^0 \in D$ and of the white-box functions u_j, $j \in \Gamma'$ on their input data $D_j := \{(x_j)_{i\in p_j}; x^0 \in D\}$. We assume

$$\text{genericity of } u_j|_{D_j}, \text{ for all } j \in W \tag{1.12}$$

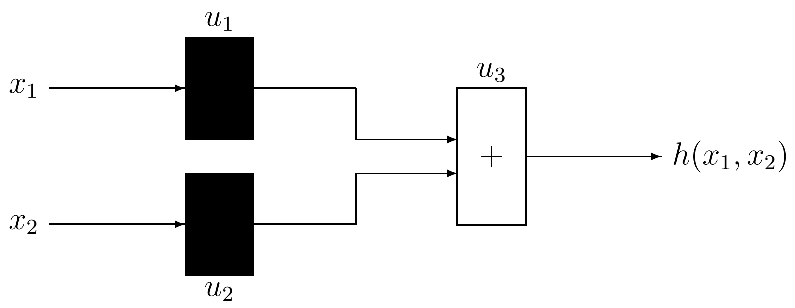

Figure 1.1: Nonunique identification due to additive white box.

in the C^ν topology.

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

$$u_1(x_1) + u_2(x_2) = h(x_1, x_2) \tag{1.13}$$

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

$$u_1(x_1) + c, \quad u_2(x_2) - c. \tag{1.14}$$

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

$$u_{j'} \circ \sigma, \quad \sigma^{-1} \circ u_j. \tag{1.15}$$

Here σ 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 (1.15) a *trivial calibration* by σ. The calibration (1.14) 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. It is worth noting that neither of these calibrations impedes the global extrapolability of the overall input-output relation h from the low-dimensional data base D to the cuboid spanned by D.

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.

We are now ready to state our main result. For a previous presentation without proof see [Sch00].

Theorem 1.1 *Consider a structured hybrid model (1.9) on a tree* Γ*. Let assumptions (1.10)–(1.12) hold for the global input-output function* h *on the data base* D*, and for the scalar white box and black-box functions* u_j*, all of class* C^ν *with* $\nu \geq 5$*. We also assume that any white box has at least two inputs.*

Then the black-box functions $u_j \in B$ *are determined uniquely on their respective input data* $(x_i)_{i \in p_j}$ *in* D_j*, up to trivial calibrations of the form (1.15).*

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

The remaining sections are organized as follows. In section 2 we show how theorem of Frobenius type reduces the identification problem to only one single elementary type of a three layer hybrid model: the *black-white-black tree*. In section 3 we complete the identification of this remaining elementary type, by local normal forms. Section 4 is devoted to the genericity assumptions on h and 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 of open problems, 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 Dimitri Turaev, for helpful discussions.

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 ω, the Frobenius theorem asserts the existence of manifolds of codimension 1 such that ω 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 $u, \tilde{u} \in C^\nu(\mathbb{R}^n, \mathbb{R})$ possess nonvanishing parallel gradients:*

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

for some positive scalar function $\alpha \in C^{\nu-1}(\mathbb{R}^n, \mathbb{R})$. Then

$$u = \sigma \circ \tilde{u}, \tag{2.2}$$

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

An analogous statement holds for negative α and σ'.

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

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

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

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

This proves the lemma. $\bowtie$

We now apply the lemma to any black box $u = u_j(x_i)_{i\in p_j}$ of the first layer, $j \in \ell_1$, with input vector $\xi := (x_i)_{i\in p_j} \in \mathbb{R}^{|p_j|}$. Without loss of generality we may indeed assume $D_j = \mathbb{R}^{|p_j|}$ for the data base D; see section 5. For a suitable C^ν-function Ψ which encodes the remaining hybrid model, and with a canonical decomposition

$$x^0 = (\xi, \eta), \tag{2.5}$$

we can write

$$\Psi(u(\xi), \eta) = h(\xi, \eta). \tag{2.6}$$

Differentiation with respect to ξ yields

$$\Psi_u \cdot u_\xi = h_\xi. \tag{2.7}$$

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 $\alpha \; inC^{\nu-1}$, as required by assumption (2.1) of lemma 2.1. Indeed,

$$\alpha(\xi) := \Psi_u(\tilde{u}(\xi), \eta)/\Psi_u(u(\xi), \eta) \tag{2.8}$$

can be written as a function of ξ alone, along a suitable $|p_j|$-dimensional submanifold of the database $s \mapsto (\xi(s), \eta(s))$, where $s \mapsto \xi(s)$ can be inverted of class C^ν. Here we use the dimension assumption (1.11) which asserts $\dim D \geq |p_j|$ for the black box $j \in B$. As a consequence, lemma 2.1 implies

$$u_j = \sigma_j \circ \tilde{u}_j, \tag{2.9}$$

i.e., uniqueness of the black box u_j up to a calibration of its values by the scalar C^ν-function σ_j.

In passing, we note that the level surfaces of u_j can be determined constructively, from the gradients (2.7), by successive solution of families of suitable ordinary differential equations; see [Arn83].

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 σ_j^{-1} into the black successor box u_j, by calibration; see (1.15). Thus u_j has been identified, up to an unavoidable trivial calibration.

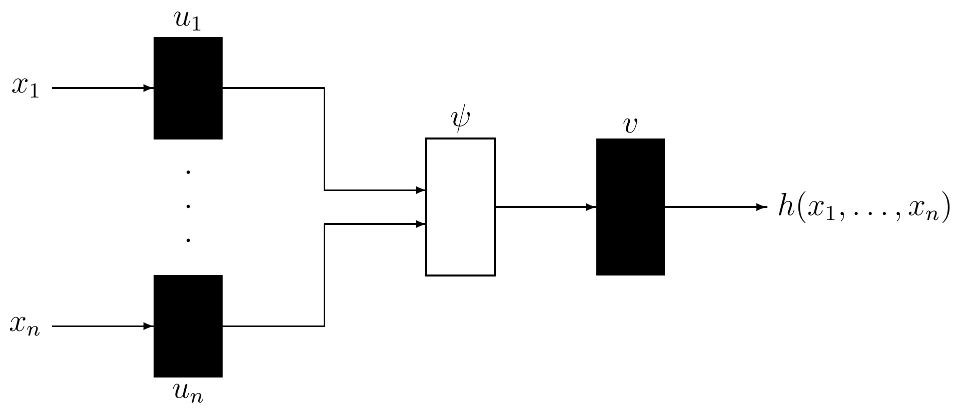

Figure 2.1: The elementary black-white-black hybrid tree.

Next suppose the successor box is white, $j' \in W$. Analogously to (2.9), and up to calibration, 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'. We thus arrive at a new hybrid model of the form

$$\Phi(u_{j'}((\sigma_j(x_j))_{j \in p_{j'}}), \eta) = h((x_j)_{j \in p_{j'}}, \eta). \tag{2.10}$$

Here we consider the variables x_j as new inputs, because they are preceded by only white boxes, and thus are uniquely determined by the original input vector x^0. As above, η denotes the remaining input components.

We now view $\eta = \eta_0$ in (2.10) as a constant, and define the scalar function

$$v := \Phi(\cdot, \eta_0) \in C^\nu(\mathbb{R}, \mathbb{R}), \tag{2.11}$$

which we consider as a new black box. With the abbreviation $h_0 := h(\cdot, \eta_0)$ we can then rewrite (2.10) as

$$v(u_{j'}((\sigma_j)_{j \in p_{j'}})) = h_0. \tag{2.12}$$

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 ψ, and reminds of the origin of the calibrations σ_j from the black boxes u_j.

Unfortunately, the 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 η_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 (2.12) as valid. In sections 3 and 4 we will therefore show how to uniquely determine all black boxes u_j, viz. the above calibrations σ_j, in this tree, from local information on the ν-jet of $\tilde{h} \in C^\nu$. This will complete the identification of all black boxes u_j 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 also complete the identification of the entire tree.

3 Identification of the elementary black-white-black tree by local normal forms

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

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

which was posed in (2.12); see also Figure 2.1. We develop a local normal form approach which, for generic white-box functions ψ, 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 trivial calibrations.

Normal forms are representatives of group orbits. The Jordan normal form of linear algebra 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. [Van89] and the references there. From the same view point, equation (3.1) states that h and ψ lie in the same group orbit under the action of the transformations by the diffeomorphisms $u_1, \ldots, u_n, v$ of the real line. For simplicity, we consider smooth functions; the obvious adaptation to C^ν-functions just limits considerations to using at most ν derivatives.

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

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

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

$$v(\psi(u_1(x_1), \ldots, u_n(x_n))) = H(x_1, \ldots, x_n), \tag{3.3}$$

where the normal form H *of* ψ *satisfies*

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

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

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

Under the normalizations

$$u_j(0) = 0, v(0) = 0, v'(0) = 1, \tag{3.6}$$

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 $n \geq 2$. Without loss of generality, we assume $\psi(0) = 0$. We use the abbreviations $x = (x_1, \ldots, x_n)$, $u(x) = (u_1(x_1), \ldots, u_n(x_n))$ and

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

Inserting $x = t\mathbf{e}_j$ in the variant

$$\psi(u(x)) = v(H(x)) \tag{3.8}$$

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

$$\begin{aligned} \psi_j(u_j(t)) &= \psi(u_j(t)\mathbf{e}_j) = \psi(u(t\mathbf{e}_j)) = v(H(u(t\mathbf{e}_j))) = \\ &= v(H(u_j(t)\mathbf{e}_j)) = v(t). \end{aligned} \tag{3.9}$$

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

$$u_j := \psi_j^{-1} v. \tag{3.10}$$

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

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

Since $\varphi(0) = 0$, $\varphi'(0) = n$, we say that (3.11) is equivalent to a conjugation of φ to its linearization, by v. By standard normal form theory for diffeomorphisms φ, 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 $t = 0$:

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

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 $J(u_j) = J(v) = J(\mathrm{id})$ for the k-jets J at 0. Here we have substituted H itself for ψ in (3.8). This reduces our previous analysis to the special case $\psi_i = \mathrm{id}$, $\varphi = n \cdot \mathrm{id}$, 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

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

By the normalization (3.6), $v'(0) = 1$, this implies that the k-jets of v and id coincide at $t = 0$. Since $u_j \equiv v$, the same holds true for all transformations u_j. This proves the lemma. $\bowtie$

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 $v'(0)$ leads to a slightly more restricted 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 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 ψ 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 λ 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 λ. Similarly $u_j(t) = \lambda t$ in (3.13). Therefore the jet of H at $x = 0$ satisfies $H(\lambda x) = \lambda H(x)$. More precisely

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

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

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.

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* $k = (k_1, \ldots, k_n)$ *let* $H_k := H^{(k)}(0)$ *denote the corresponding partial derivative of the normal form* H *at zero. Then the normal*

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 $|k| \le p$ satisfying the normal form constraints (3.20)–(3.23).

form conditions (3.4), (3.5), (3.16) are satisfied if, and only if

$$H_k = 0, \qquad \text{for } k = p \cdot e_j \tag{3.20}$$

$$H_k = 1, \qquad \text{for } k = e_j \tag{3.21}$$

$$H_{k_0} = 1 \tag{3.22}$$

$$\sum_{|k|=p} \frac{1}{k!} H_k = 0 \tag{3.23}$$

hold for all orders $p \pm 1$ and for all $j = 1, \dots, n$. The even order $|k_0|$ can be chosen to be 4, for $n = 2$. For $n \ge 3$, we can choose $|k_0| = 2$. Of course $k_0 \ne pe_j$ must not contradict (3.20). For the resulting dimensions of the spaces of H_k up to fixed order p see table 3.1.

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

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

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

for $n \ge 3$. For $n = 2$ and $p = 2, 3$, the numbers in (3.24) have to be increased by 1. For the resulting dimensions $d(n,p)$ see Table 3.1.

Proof: The proof follows by elementary counting. $\bowtie$

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*

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

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

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

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 ν-jet of the normal form H, evaluated at $x = 0$. More generally, define,

$$I(\psi, u^0) := I(H, 0) \tag{3.27}$$

for functions ψ 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:

$$v(\psi(u(x))) = h(x). \tag{3.28}$$

In other words, the given white-box function ψ 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

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

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

Theorem 3.6 *Let $I(H, 0)$ be any function of the ν-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 ψ, and with global input-output data h; all with nonvanishing (partial) derivatives. Then*

$$I(\psi, u^0) = I(h, x^0) \tag{3.30}$$

holds at $u^0 := u(x^0) = (u_1(x^0), \ldots, u_n(x^0))$*, 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 ψ, 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:

$$u^0 = u(x^0). \tag{3.31}$$

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), (3.31) will be addressed in section 4.

We conclude this section with several examples. We first compute the group invariants

$$I_k(\psi, u^0) = H_k \tag{3.32}$$

for $n \geq 4$ inputs and order $p = |k| \leq 2$. Deviating from the fourth order normalization (3.15), (3.22), we consider the second order normalization

$$H_{x_1 x_2} = 1. \tag{3.33}$$

See the remark following the proof of Corollary 3.2. This is also suggested by Table 3.1, $d(n,p) \geq 4$, for $n \geq 4$ and $p = 2$. With the variant normal form (3.4), (3.5), (3.32) we thus expect

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

invariants, in terms of the Taylor coefficients

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

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

We first show how the normal form transformations

$$v(\psi(u(x))) = H(x) \tag{3.36}$$

are determined, up to second order, by the normal form constraints (3.19), (3.20), (3.22), (3.32). Differentiating (3.36) with respect to x_j, at $x = 0$, we obtain

$$1 = H_{x_j} = v'\psi_j u'_j, \tag{3.37}$$

for all j. Differentiating twice provides

$$H_{x_{j_1}x_{j_2}} = v''\psi_{j_1}u'_{j_1} \cdot \psi_{j_2}u'_{j_2} + v'\psi_{j_1j_2}u'_{j_1}u'_{j_2} + \delta_{j_1j_2}\psi_{j_1}u''_{j_1}, \tag{3.38}$$

where $\delta_{j_1j_2}$ is the Kronecker symbol. Expressing u'_j by the normal form constraint (3.37) and evaluating (3.38) for $j_1 \neq j_2$ we obtain

$$H_{j_1j_2} = \frac{1}{v'}\left(\frac{v''}{v'} + \frac{\psi_{j_1j_2}}{\psi_{j_1}\psi_{j_2}}\right). \tag{3.39}$$

By summation over all $n(n-1)$ pairs $j_1 \neq j_2$, the normal form constraint $\sum H_{j_1j_2} = 0$ of (3.22) implies

$$0 = n(n-1)\frac{v''}{v'} + \sum_{j_1\neq j_2} q_{j_1j_2}, \qquad \text{i.e.} \tag{3.40}$$

$$\frac{v''}{v'} = -\bar{q}. \tag{3.41}$$

Here we have abbreviated

$$\begin{aligned} q_{j_1j_2} &:= \frac{\psi_{j_1j_2}}{\psi_{j_1}\psi_{j_2}} \\ \bar{q} &:= \frac{1}{n(n-1)}\sum_{j_1\neq j_2} q_{j_1j_2}. \end{aligned} \tag{3.42}$$

To determine v' itself, we invoke the normal form condition (3.32) which, in view of (3.39), (3.41), implies

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

We therefore obtain the *invariants*

$$I_{j_1j_2}(\psi, u^0) := H_{x_{j_1}x_{j_2}} = \frac{q_{j_1j_2} - \bar{q}}{q_{12} - \bar{q}}, \tag{3.44}$$

with $q_{j_1j_2}$, $\bar{q}$ defined in (3.42). Clearly $\sum_{j_1\neq j_2} I_{j_1j_2} = 0$ and $I_{12} = 1$ are normalized, and $I_{j_1j_2} = I_{j_2j_1}$. This leaves precisely $\tilde{d}(n,p) = \frac{1}{2}n(n-1)-2$ independent invariants of order $|k| \leq 2$ in ψ, as predicted in (3.34). Moreover, the equations

$$I_{j_1j_2}(\psi, u^0) = I_{j_1j_2}(h, x^0), \tag{3.45}$$

which hold by theorem 3.5, 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 $\tilde{d}(n,p)$ remaining invariants, because these transformations are defined uniquely by ψ itself. Moreover the order of the required ψ-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, relations (3.20), (3.23) for $p = 2$ imply

$$H_{xy} + H_{yz} + H_{xz} = 0 \tag{3.46}$$

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

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

We recall that ψ-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 ψ-derivatives of (3.47). Of course the relation (3.46) is now obvious from the explicit form (3.47). Picking $k_0 = (1, 0, 1)$ from the normalizations (3.16), (3.22) of corollaries 3.2, 3.3 obviously eliminates $v'(0)$. We may then choose

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

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 ψ 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.

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

Here the summation runs over all permutations of the subscript indices $\{1, 2, 3\}$. Subscripts again indicate partial derivatives of ψ. Similarly

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

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 $H_{xy} = 1$, and in view of relations (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):

$$H_{xxy} - H_{xyy} = \frac{v'(0)^2}{2\psi_1^3\psi_2^3} \left(\psi_1^2\psi_{12}\psi_{22} + \psi_1\psi_2^2\psi_{122} - \{12\}\right) \tag{3.51}$$

together with its two other substitution associates. Here the term $\{12\}$ repeats all other terms inside the parenthesis with derivative indices $1, 2$ of ψ 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 other term with $1, 2$ interchanged. The obvious antisymmetry follows from (3.20), (3.23). At order $p = 4$, we obtain

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

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

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

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 ψ and the given data h for the n unknown black-box functions, $u_1, \ldots, u_n$. 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.

4 Genericity

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

$$I_K(\psi, u^0) = I_K(h, x^0) = H_K \tag{4.1}$$

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

$$u^0 = u^0(x^0) \tag{4.2}$$

of system (4.1), for given global input-output data h, and for generic white-box functions $\psi \in X = BC^\nu(\mathbb{R}^n, \mathbb{R})$. Here the open subset X is defined by the condition of nonzero partial derivatives ψ_{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 ψ in a residual subset of $X = BC^\nu$, viz. at least for ψ in a countable intersection of open dense sets. See [Hir76] and in particular [AR67] for a background.

To prove our claim, we show that the map

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

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

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

Here $\tilde{X} \subseteq X$ is an open dense subset, and Δ 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 zero values 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

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

for generic white-box functions $\psi \in \tilde{X}$. But the partial derivative of $I(\psi, \cdot, \cdot)$ cannot be surjective, for any fixed ψ, 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.

This completes the proof of theorem 1.1 for the black-white-black trees of section 3, up to the following lemma.

Lemma 4.1 *Choose $\nu \geq \nu(n)$ large enough. Then there exist index sets $K \subseteq \{2, \ldots, \nu\}^n$ of cardinality $2n+1$ and an open dense subset $\tilde{X} \subseteq X = BC^\nu$, such that the map $\mathcal{I}$ defined in (4.4) possesses a surjective partial derivative with respect to ψ, 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_ψ. By the group action (3.1) of the transformations v, u_j, we reduce surjectivity of I_ψ to the special case where ψ is already in normal form: $\psi = H$. By linearity of the space of normal forms in the jet space, surjectivity at $\psi = H$ is trivial, for proper choices of the index set K and of the open dense subset $\tilde{X} \subset X$ of white-box functions ψ.

Concerning differentiability, we first observe that the unique explicit $\nu-$jets of the transformations u, v to normal form H depend C^1 on $\psi \in C^\nu$; 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 φ by $\varphi(v(t)) = v(nt)$, this follows from the explicit representation of $v(t) = t + \tilde{v}(t)$ by the convergent sum

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

Here we have also decomposed $\varphi(t) = nt + \tilde{\varphi}(t)$, and we denote the backwards iteration of φ 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_ψ at any $\psi = \psi^0 \in \tilde{X}$, $u = u^0 \in \mathbb{R}^n$. Then

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

is also surjective at any $\psi = \psi^0 \in \tilde{X}$ and 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 ψ 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 ψ in the normal form space. If 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. $\bowtie$

5 Proof of theorem 1.1

In this section we summarize the proof of theorem 1.1, as given in sections 2-4. We first give a precise version of our "general position" assumption (1.11) on D. In particular we show that input data sets D_j of box j form rectangles, for black boxes $j \in B$. We then recall the reduction of the identification problem to black-white-black trees, of section 2, as well as the identification in section 3 for the generic intermediate white boxes $\psi = u_j$, $j \in W$, of section 4. Together, this will prove theorem 1.1.

To address assumption (1.11) on general position of D, 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_ι 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 i belonging to the same layer ℓ_k, and the sets J_i decompose the input layer ℓ_0. By monotonicity assumption (1.10) and the chain rule, the partial derivatives of x_i with respect to any input x_ι, $\iota \in J_i$, are nonzero. Because the J_i are moreover disjoint, for inputs $i \in p_j$ of any box j, the function

$$(x_i)_{i \in p_j} \tag{5.1}$$

of the input vector $x^0 \in \mathrm{I\!R}^m$ possesses only regular values. For every black box $j \in B$, the function values $(x_i)_{i\in p_j} \in D_j$ therefore define a foliation of the input space, with leaves of codimension $|\, p_j\,|$. We assume specifically, that the data base D of dimension $d = \max_{j\in B} |p_j|$ is transverse to these foliations, for any black box $j \in B$. Moreover, we assume transversality of D to the foliation of codimension 1 defined by any single input variable x_i, $i < N$. Genericity of such assumptions is standard [AR67].

As a consequence of this general position assumption on D, the values of any x_i form open intervals, for inputs $x^0 \in D \subset \mathrm{I\!R}^m$. Similarly, the input data sets $D_j \subset \mathrm{I\!R}^{|p_j|}$ of black boxes u_j form open rectangles. For recursions in the spirit of sections 2–3 we may always rescale these rectangles to coincide with $\mathrm{I\!R}^{|p_j|}$, if necessary.

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 ℓ_z. 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).

From table 3.1 we conclude that the maximal order ν of derivatives necessary to capture $2n+1$ independent invariants H_K is $\nu = 5$, which curiously arises at white boxes ψ 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 rescale all input intervals x_j, $j \in \ell_1$, to all of $\mathrm{I\!R}$, replacing in particular $D_{ji} \subseteq \mathrm{I\!R}^{|p_{ji}|}$ by $\mathrm{I\!R}^{|p_{ji}|}$ for all black boxes $j' \in \ell_2$. We can then repeat the same identification process for layer ℓ_2 and so on, inductively, through the scalar hybrid tree. This proves theorem 1.1.

6 Industrial application

As an example for an industrial application an SHM model for a continuous polymerization plant will be described. 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 [enc98]. The material produced in the process under consideration had special material structures, as were 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 was 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 parameter had caused the training of pure black-box models to fail. Therefore optimization and model based control of MI could only be performed using an SHM model for the reactor.

Using the chemical know-how about the reactive part, an SHM model 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 was not established with the necessary reliability. Therefore a black-box model had to be used for the dependence w of MI on the average chain length c_L:

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

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 per mole of monomers:

$$c_L = \psi_3(Y, N_c) \tag{6.2}$$

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

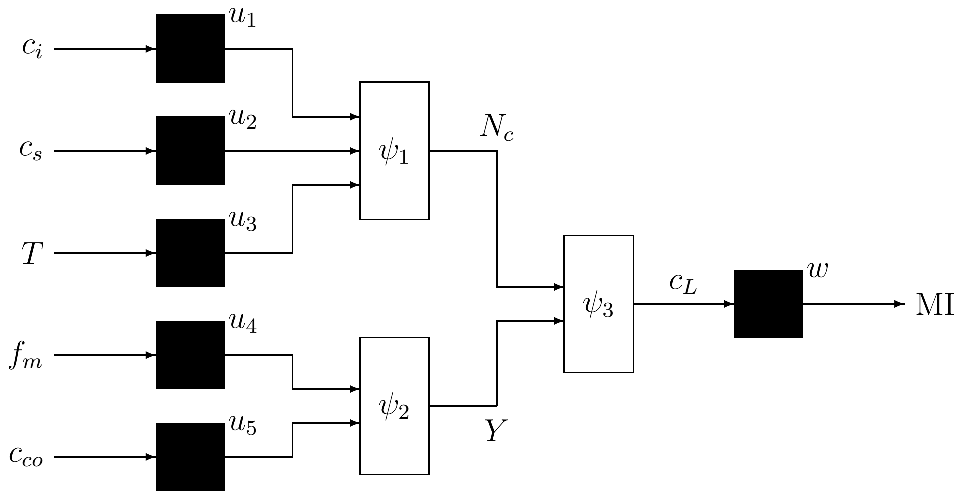

Figure 6.1: SHM for the polymerization mechanism. The vector $c = (c_i, c_s, c_{co}) \in \mathrm{IR}^3$ describes the concentrations of initiator, chain length controller, and comonomer species; f_m denotes the monomer flow and T the feed flow temperature.

and N_c are input variables for a white-box subprocess ψ_3 in which the average chain length 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 sub-processes 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:

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

with unknown black-box functions $u_1, \dots, u_5$ and w depending on only one input variable each. The processes ψ_1 and ψ_2 have been modeled by known white boxes. Moreover, the input variable c_L to the melt-index black box $\mathrm{MI} = w(c_L)$ has been replaced by the white-box output $c_L = \psi_3(N_c, Y)$. Alternatively, the three white boxes ψ_1, ψ_2, ψ_3 could have been combined into a single white box ψ 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:

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

the dimension of the required data-base has been reduced from five to one. This reduces the data demand significantly. Even more importantly, the SHM model could be used for extrapolations even when it had been trained on highly correlated data sets, only. The model was 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, had been trained on the same data. For the underlying real process, the neural network alone required about 2000 data sets to predict the melt index properly, whereas the SHM could be used for melt-index prediction after training on only 200 data sets. Moreover, the prediction quality of the SHM was about five percent higher than that of the neural network, although the training set had been 10-fold smaller. Moreover, due to the low dimension of the data base, the SHM allowed the analysis of the impact of monomer impurities on the melt index from plant data without any additional experiments.

7 Summary and discussion

We have considered trees $j \in \{1, \ldots, N\}$ 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 data base D of dimension $d \geq \max_{j \in B} |p_j|$, 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 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 from coefficients H_K, for the black-white-black tree of section 3, and the resulting grossly overdetermined systems

$$I_K(\psi, u^0) = I_K(h, x^0) \tag{7.1}$$

already provide many consistency constraints to be satisfied by 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, say of cardinality n, may lead to different solutions $u^0_K(x^0; \psi, h)$ and different numerical condition numbers for the solutions of (7.1). Dependence of the solution on K indicates inconsistency of data.

The high differentiability requirements on ψ, 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 three layer hybrid model

$$\psi(u_1(x_1), u_2(x_2)) = h(x_1, x_2) \tag{7.2}$$

with nonmonotone cubic white box ψ. We have encountered up to 11 solutions u^0_j of the associated invariant system

$$\begin{aligned} \psi(u^0_1, u^0_2) &= h^0_1 \\ \frac{\psi_{u_1 u_2}}{\psi_{u_1}\psi_{u_2}}(u^0_1, u^0_2) &= h^0_2. \end{aligned} \tag{7.3}$$

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 Γ. This case, so particularly relevant in chemical industry, 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 used for control as well as for optimization applications.

(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

[AR67] R. Abraham and J. Robbin. *Transversal Mappings and Flows*. Benjamin Inc., Amsterdam, 1967.

[Arn63] V. I. Arnol'd. On functions of three variables. *AMS Transl.*, **28**:51–54, (1963).

[Arn83] V. I. Arnol'd. *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer-Verlag, New York, 1983.

[BvCV98] H. A. B. Te Braake, H. J. L. van Can, and H. B. Verbruggen. Semi-mechanistic modelling of chemical processes with neural networks. *Eng. Appl. Art. Intell.*, **11**:507–515, (1998).

[DSL95] M. Dors, R. Simutis, and A. Lübbert. Hybrid process modelling for advanced process state estimation, prediction and control examplified in a production-scale mammalian cell culture. *ACS Symposium Series*, **613**:144–154, (1995).

[enc98] *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley, 1998.

[Fri67] 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).

[Hil00] D. Hilbert. Mathematische Probleme. *Archiv der Mathematik und Physik*, **1**:44–63, (1900).

[Hir76] M. W. Hirsch. *Differential Topology*. Springer-Verlag, New York, 1976.

[Kol63] 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).

[PU92] D. C. Psichiogos and L. H. Ungar. A hybrid neural network — first principle approach to process modelling. *AIChE Journal*, **38**:1499–1511, (1992).

[Sch96] A. Schuppert. New approaches to data-oriented reaction modelling. In *Proceedings of the 3rd Workshop on Modelling of Chemical Reaction Sytems*. Springer-Verlag, Heidelberg, 1996.

[Sch00] A. Schuppert. Extrapolability of structured hybrid models: A key to optimziation of complex processes. In B. Fiedler et al., editor, *International Conference on Differential Equations*, pages 1135–1151. World Scientific Publ., Singapore, 2000.

[TK94] M. L. Thompson and M. A. Kramer. Modelling chemical processes using prior knowledge and neural networks. *AIChE Journal*, **40**:1328–1340, (1994).

[Van89] A. Vanderbauwhede. Centre manifolds, normal forms and elementary bifurcations. In U. Kirchgraber and H. O. Walther, editors, *Dynamics reported*, volume 2, pages 89–169. Teubner & Wiley, Stuttgart, 1989.

[Vit54] A. G. Vitushkin. On Hilbert's thirteenth problem. *Dokl. Akad. Nauk SSSR*, **95**:701–704, (1954).

[Vit77] A. G. Vitushkin. On representation of functions by means of superpositions and related topics. *L'Enseignement Mathematique Ser. II.*, **23**:255–320, (1977).