Published paper
Anderson, S.R. and Kadirkamanathan, V. (2007) Modelling and identification of non-linear deterministic systems in the delta-domain, Automatica**,** Volume 43 (11), 1859 - 1868.

Modelling and identification of non-linear deterministic systems in the delta-domain

S. R. Anderson [a,*], V. Kadirkamanathan [b]

[a] *Neural Algorithms Research Group, Department of Psychology, University of Sheffield, Western Bank, Sheffield, S10 2TP, UK*

[b] *Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK*

Abstract

This paper provides a formulation for using the delta-operator in the modelling of non-linear systems. It is shown that a unique representation of a deterministic non-linear auto-regressive with exogenous input (NARX) model can be obtained for polynomial basis functions using the delta-operator and expressions are derived to convert between the shift- and delta-domain. A delta-NARX model is applied to the identification of a test problem (a Van-der-Pol oscillator): a comparison is made with the standard shift operator non-linear model and it is demonstrated that the delta-domain approach improves the numerical properties of structure detection, leads to a parsimonious description and provides a model that is closely linked to the continuous-time non-linear system in terms of both parameters and structure.

Key words: Delta operator, Nonlinear system identification, NARX model

1 Introduction

Middleton and Goodwin (1986) have renewed interest in the use of a gradient based discrete-time operator, δ to parameterise data-driven system models. This has been shown to improve on the numerical ill-conditioning problems found when using the absolute valued forward shift operator q (Li and Fan, 1997), especially under conditions of fast-sampling.

The δ-operator has been widely investigated in problems in the areas of adaptive signal processing (Fan and De, 2001), systems' modelling (Kuznetsov et al., 1999; Fan et al., 1999; Larsson et al., 2006) and control (Middleton and Goodwin, 1990; Lauritsen et al., 1997; Suchomski, 2003; Shim and Sawan, 2006). In the context of non-linear systems investigations into the use of the δ-operator are currently limited: analysis of non-linear δ-domain models (in terms of generalised frequency response functions) was conducted in Chadwick et al. (2006) and a sampled data model for non-linear continuous-time systems has been developed in Yuz and Goodwin (2005).

* Corresponding author.
Email address: `s.anderson@sheffield.ac.uk` (S. R. Anderson).

An advantage of modelling linear systems in the δ-domain is that, whilst it provides an exact discrete-time representation of the system, the identified model has structural similarity to the continuous-time differential equation describing system dynamics; additionally the parameters of the identified model approach the continuous time values as the sample time tends to zero (Soderstrom et al., 1997; Fan et al., 1999). Such a direct equivalence between continuous and discrete-time non-linear system descriptions does not necessarily exist (Monaco and Normand-Cyrot, 1995). However, it may be expected in some cases that the use of the δ-operator can lead to the identification of discrete-time non-linear models that retain a link to the continuous-time system description, in terms of both structure and parameters. This aspect is investigated here along with the numerical improvements that result from use of the δ-operator in non-linear system identification.

The non-linear auto-regressive moving average with exogenous inputs (NARMAX) model (Leontaritis and Billings, 1985; Liu et al., 2000) is able to represent a wide range of non-linear dynamical systems, via the use of a range of potential basis functions. This paper takes the approach of defining a deterministic δ-NARX model and investigating its relationship to the corresponding q-domain non-linear model for the case of polynomial basis functions. An advantage of using polynomial basis functions is that it is often more straightforward to find some relationship between the model and the physical system.

An investigation into the effect of sample time choice on non-linear system identification was carried out in Billings and Aguirre (1995). One conclusion was that fast-sampling hampers structure selection. This is due to the numerical similarity between samples. The advantage of utilising the δ-operator is that it naturally overcomes the problem of numerical similarity, especially in conditions of fast-sampling. Therefore simply describing the model in the δ-domain directly results in the numerical properties of structure detection.

There is an exact relationship between the polynomial δ- and q-domain NARX models, which is shown in section 3. This theoretical equivalence, which does not necessarily exist for all classes of non-linear model, implies that the use of either domain should lead to equivalent models in the practical application of non-linear system identification techniques. Any differences in both model parsimony and consistency of non-linear term selection should be ascribed to the procedure of structure detection and parameter estimation, which are affected by numerical properties. In order to elucidate these differences the properties of the δ-operator are investigated through an example problem, with regard to: convergence of the discrete-time model representation to the continuous-time system, consistency of structure detection and model parsimony.

The structure of the paper is as follows: section 2 provides background information on the δ-operator and the polynomial NARX model, and presents the δ-domain polynomial model. Section 3 explores the link between the polynomial δ-NARX model and the equivalent q-domain model. Section 4 compares from a theoretical perspective the numerical problems that would be expected when using the q-operator and the improvements due to use of the δ-operator. Section 5 demonstrates non-linear system identification in both the q- and δ-domains, on the test problem of the identification of a Van-der-Pol oscillator. Finally the main results of the paper are summarised in section 6.

2 Background

A discrete-time single-input single-output non-linear system can be described by the deterministic NARX model

$$y(t+n) = f\left[y(t), \ldots, y(t+n-1), u(t), \ldots, u(t+n-1)\right] \quad (1)$$

where $f(.)$ is a non-linear function and $u(t)$ and $y(t)$ are sampled input and output data respectively. The structure of $f(.)$ is usually unknown unless certain *a priori* knowledge is available.

In order to consider a general NARX model term irrespective of categorising the input or output term explicitly, the operator $\phi(t)$ is defined, which may be either an input or output term,

$$\phi(t) \in \{u(t), y(t)\} \quad (2)$$

where the signal $\phi(t)$ is defined by the implementation of the system model. This allows model terms to be specified via a numerical indexing notation rather than an alphabetical system, which will prove useful later when mapping between q- and δ-domain descriptions. Using this notation the general polynomial NARX term $\chi(t)$ is defined as

$$\chi(t) = \prod_{j=1}^{p} q^{n_j} \phi_j(t) \quad (3)$$

where p is the number of cross-product terms (commonly known as the order of non-linearity), n_j is the forward shift delay of the j^{th} cross-product term and q is the forward shift operator, that is $qu(t) = u(t+1)$.

Definition 2.1 *The expansion of a NARX model corresponding to (1) in the form of polynomial basis functions is*

$$q^n y(t) = \sum_{l=1}^{w} \theta_l \chi_l(t) \quad (4)$$

where

$$\chi_l(t) = \prod_{j=1}^{p_l} q^{n_{l,j}} \phi_{l,j}(t) \quad (5)$$

and w is the number of model terms within the NARX model, p_l is the number of cross-product terms within the l^{th} model term (known as the order of non-linearity), $n_{l,j}$ is the forward shift delay of the l^{th} model term and j^{th} cross-product term and θ_l are model parameters.

Similarly to the q-domain NARX model, the deterministic δ-NARX model is defined as

$$\delta^n y(t) = f\left[y(t), \delta y(t) \ldots, \delta^{n-1} y(t), u(t), \delta u(t), \ldots, \delta^{n-1} u(t)\right] \quad (6)$$

where

$$\delta = \frac{q-1}{T} \quad (7)$$

where T is the interval between samples. The δ-operator has the useful property, given a differentiable signal $y(t)$ that

$$\lim_{T \to 0} \delta y(t) = \frac{d}{dt} y(t) \quad (8)$$

The δ-domain polynomial basis function ψ can be defined similarly to the q-domain as

$$\psi(t) = \prod_{j=1}^{p} \delta^{n_j} \phi_j(t). \tag{9}$$

Definition 2.2 *The expansion of the δ-NARX model corresponding to (6) in the form of polynomial basis functions (using the previously defined notation) is*

$$\delta^n y(t) = \sum_{l=1}^{w} \zeta_l \psi_l(t) \tag{10}$$

where

$$\psi_l(t) = \prod_{j=1}^{p_l} \delta^{n_{l,j}} \phi_{l,j}(t) \tag{11}$$

and ζ_l are the δ-domain model parameters.

3 Relationship between the q- and δ-domain polynomial NARX models

The equivalence of the model in each domain (δ and q) is important for translation and interpretation, which has already been established for linear models (Neuman, 1993). This section establishes a relationship between the δ- and q-domain polynomial NARX models.

Lemma 3.1 *The expression that maps from a single l^{th} q-domain NARX model term $\chi(t)$ defined in (3) to the δ-domain is given by*

$$\prod_{j=1}^{p} q^{n_j} \phi_j(t) = \sum_{k=1}^{r} T^{\bar{m}_k} \prod_{j=1}^{p} c_{m_{k,j},j} \delta^{m_{k,j}} \phi_j(t) \tag{12}$$

where

$$r = \prod_{j=1}^{p} (n_j + 1) \tag{13}$$

$$\bar{m}_k = \sum_{j=1}^{p} m_{k,j} \tag{14}$$

$$m_{k,j} = \mod \left(\left\lfloor \frac{k-1}{\prod_{i=1}^{j-1} n_i + 1} \right\rfloor, n_j + 1 \right) \tag{15}$$

$$c_{m_{k,j},j} = \frac{n_j!}{(n_j - m_{k,j})! m_{k,j}!} \tag{16}$$

where the floor function $\lfloor x \rfloor$ and modulus function mod (x) *are defined as*

$$\lfloor x \rfloor = y \quad \textit{where} \quad x - 1 < y \leq x, \quad y \in \mathbb{Z}, \quad x \in \mathbb{R} \tag{17}$$

$$\mod (x, y) = x - \left(y \times \left\lfloor \frac{x}{y} \right\rfloor \right) \tag{18}$$

where $\mathbb{Z}$ is the set of integer numbers and $\mathbb{R}$ is the set of real numbers.

Lemma 3.2 *The expression that maps from a single δ-domain NARX model term $\psi(t)$ defined in (9) to the q-domain is given by*

$$\prod_{j=1}^{p} \delta^{n_j} \phi_j(t) = \sum_{k=1}^{r} T^{-\bar{m}_k} \prod_{j=1}^{p} c^-_{m_{k,j},j} q^{m_{k,j}} \phi_j(t) \tag{19}$$

where

$$c^-_{m_{k,j},j} = (-1)^{m_{k,j}} \frac{n_j!}{(n_j - m_{k,j})! m_{k,j}!} \tag{20}$$

Theorem 3.1 *The mapping of a δ-domain to q-domain polynomial NARX model is unique, preserves the order of non-linearity and preserves the input/output term order. The mapping of a full model of w terms is*

$$q^n y(t) = \sum_{l=1}^{w} \theta_l \sum_{k=1}^{r_l} T^{\bar{m}_{l,k}} \prod_{j=1}^{p_l} c_{l,m_{l,k,j},j} \delta^{m_{l,k,j}} \phi_{l,j}(t) \tag{21}$$

where

$$r_l = \prod_{j=1}^{p_l} (n_{l,j} + 1) \tag{22}$$

$$\bar{m}_{l,k} = \sum_{j=1}^{p_l} m_{l,k,j} \tag{23}$$

$$m_{l,k,j} = \mod \left(\left\lfloor \frac{k-1}{\prod_{i=1}^{j-1} n_{l,i} + 1} \right\rfloor, n_{l,j} + 1 \right) \tag{24}$$

$$c_{l,m_{l,k,j},j} = \frac{n_{l,j}!}{(n_{l,j} - m_{l,k,j})! m_{l,k,j}!} \tag{25}$$

Theorem 3.2 *The mapping of a q-domain to δ-domain polynomial NARX model is unique, preserves the order of non-linearity and preserves the input/output term order. The mapping of a full model of w terms is*

$$\delta^n y(t) = \sum_{l=1}^{w} \zeta_l \sum_{k=1}^{r_l} T^{-\bar{m}_{l,k}} \prod_{j=1}^{p_l} c^-_{l,m_{l,k,j},j} q^{m_{l,k,j}} \phi_{l,j}(t) \tag{26}$$

where r_l, $\bar{m}_{l,k}$ and $m_{l,k,j}$ are as defined in (22), (23) and (24), and

$$c^-_{l,m_{l,k,j},j} = (-1)^{m_{l,k,j}} \frac{n_{l,j}!}{(n_{l,j} - m_{l,k,j})! m_{l,k,j}!}. \tag{27}$$

Remark 3.1 *It has been shown that the mapping of each term from one domain (either q or δ) induces many more terms in the target domain. Hence, the mapped model will not be a useful vehicle for system interpretation and* *analysis due to the large number of terms resulting from mapping all but the most simple of models. This suggests that the identification should be performed in the appropriate domain, whether q or δ. This will lead to parsimonious model descriptions useful for modelling, analysis and control.*

Remark 3.2 *The resultant expressions for mapping between model domains are the same for NARMAX models, when the model is augmented by error terms.*

4 Identification of the non-linear model

4.1 Parameter Estimation

The identification of a polynomial NARX model can be structured as a linear regression, using a predictive model, which in the δ-domain is (for a single-input single output description),

$$\delta^n y_t = \boldsymbol{\psi}_t \boldsymbol{\zeta} + \epsilon_t \tag{28}$$

where ϕ_t is the regression matrix comprised of input and output cross-product terms, $\boldsymbol{\zeta}$ is the set of parameters to be estimated and $\boldsymbol{\epsilon}$ is the modelling error, that is

$$\boldsymbol{\psi}_t = \left[\, \psi_1(t) \, \ldots \, \psi_w(t) \,\right], \tag{29}$$

$$\boldsymbol{\zeta} = \left[\, \zeta_1 \, \ldots \, \zeta_n \,\right]^T. \tag{30}$$

The prediction model for the q-domain non-linear model is similarly defined as

$$y_t = \boldsymbol{\chi}_t \boldsymbol{\theta} + \eta_t \tag{31}$$

where η_t is the model error and

$$\boldsymbol{\chi}_t = \left[\, \chi_1(t) \, \ldots \, \chi_w(t) \,\right], \tag{32}$$

$$\boldsymbol{\zeta} = \left[\, \theta_1 \, \ldots \, \theta_n \,\right]^T. \tag{33}$$

It is straightforward to show that the least-squares estimate of the parameter vectors $\boldsymbol{\zeta}^*$ and $\boldsymbol{\theta}^*$ pertaining to the δ- and q-domain models respectively, is given by

$$\boldsymbol{\zeta}^* = \left(\frac{1}{N} \sum_{t=1}^{N} \boldsymbol{\psi}_t^T \boldsymbol{\psi}_t \right)^{-1} \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{\psi}_t^T \delta^n y_t \tag{34}$$

$$\boldsymbol{\theta}^* = \left(\frac{1}{N} \sum_{t=1}^{N} \boldsymbol{\chi}_t^T \boldsymbol{\chi}_t \right)^{-1} \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{\chi}_t^T y_t \tag{35}$$

It has been demonstrated for the case of q-domain linear models that the so called information matrix, which is equivalent to the non-linear q-domain term $\frac{1}{N}\sum_{t=1}^{N} \phi_t^T \phi_t$, tends to a singular matrix as $T \to 0$ (Goodwin et al., 1992). This leads to numerical problems in the estimation of model parameters. In contrast, the linear term corresponding to $\frac{1}{N}\sum_{t=1}^{N} \psi_t^T \psi_t$ tends to the continuous-time result as $T \to 0$, which is also the case for the non-linear δ-domain model.

The numerical ill-conditioning arises for the case of the non-linear q-domain model, because the regressor terms are formed analogously to the linear case; indeed the linear terms are included as a subset of the polynomial q-domain non-linear model.

4.2 Model Term Selection

The regression matrix corresponding to the prediction model defined in (28) can be decomposed, using for example the modified Gram-Schmidt method (Chen et al., 1989); this leads to the expression of a new prediction model where the regression matrix has orthogonal columns, which allows the independent assessment of the significance of model terms,

$$\mathbf{y} = W\mathbf{g} + \boldsymbol{\epsilon} \tag{36}$$

where $W \in \mathbb{R}^{N \times w}$ is the new regression matrix (with orthogonal columns) and $\mathbf{g}$ is the corresponding parameter vector to be estimated, that is,

$$W = \Psi A^{-1}, \tag{37}$$

$$\mathbf{g} = A\boldsymbol{\zeta}, \tag{38}$$

$$\mathbf{y} = \left[\, \delta^n y_1 \, \ldots \, \delta^n y_N \,\right]^T, \tag{39}$$

$$\Psi = \left[\, \boldsymbol{\psi}_1^T \, \ldots \, \boldsymbol{\psi}_N^T \,\right]^T, \tag{40}$$

and $A \in \mathbb{R}^{w \times w}$ is an upper triangular matrix.

The forward regression orthogonal (FRO) algorithm (Chen et al., 1989) is used to select the model structure by iteratively comparing and ranking terms by their significance; a term's significance is measured by its contribution to the variance of the target data (based on a one-step-ahead prediction in time). This metric is called the error reduction ratio (*Err*), where

$$Err_k = \frac{g_k^2 \mathbf{w}_k^T \mathbf{w}_k}{\mathbf{y}^T \mathbf{y}} \tag{41}$$

where $\mathbf{w}_k$ is the k^{th} column in W and the subscript $k \in \{1, \ldots, w\}$ denotes the k^{th} model term. Terms are picked in order of the size of their *Err*, where the largest *Err* is most significant.

A particular problem associated with the FRO algorithm occurs under conditions of fast-sampling when identifying q-domain NARX models (Billings and Aguirre, 1995); consider the Err for the case where there is only a single regressor

$$Err_1 = \frac{g_1^2 \sum_{t=1}^{N} y(t-1)^2}{\sum_{t=1}^{N} y(t)^2} \tag{42}$$

where $g_1 = \frac{\sum_{t=1}^{N} y(t)y(t-1)}{\sum_{k=1}^{N} y(t-1)^2}$. Now considering the limit as $T \to 0$ in (42):

$$\lim_{T \to 0} Err_1 = 1 \tag{43}$$

Therefore it is apparent that in the limit $T \to 0$, only one term is necessary to explain the target data, specifically $y(t-1)$. Under conditions of fast-sampling this means that $y(t-1)$ will be selected as the first significant term, overly dominating the significance of all other terms.

Discretising the data via the δ-operator naturally overcomes this problem, because columns in the regression matrix are not numerically similar. Consider the Err of a model with a single regressor, similarly to (42), but now in the δ-domain

$$Err_1 = \frac{g_1^2 \sum_{t=1}^{N} y(t)^2}{\sum_{t=1}^{N} \delta y(t)^2} \tag{44}$$

In the limit as the sample-time tends to zero

$$\lim_{T \to 0} Err_1 = \frac{g_1^2 \sum_{t=1}^{N} y(t)^2}{\sum_{t=1}^{N} \frac{dy}{dt}^2} \tag{45}$$

Evidence of these improved properties of structure selection will be seen in the next section, where the method is applied to the identification of a Van-der-Pol oscillator.

5 Identification of a Van-der-Pol oscillator

This section investigates the δ-operator modelling framework applied to a test system. The motivation for this is to show the advantages of working in the δ-domain in comparison to the q-domain, in two specific areas: (i) insight into the physical system and (ii) numerical improvements of the identification procedure. The chosen test system was a Van-der-Pol oscillator (VDPO).

5.1 Data Generation

The particular Van-der-Pol oscillator utilised in this investigation was

$$\frac{d^2}{dt^2} y(t) = 0.2 \left[1 - y^2(t)\right] \frac{d}{dt} y(t) - y(t) + u(t) \tag{46}$$

The bandwidth of the VDPO (neglecting the non-linear term) was known to be 1Hz. Therefore a conservative excitation signal was applied using a sum-of-sinusoid signal, with frequencies, ω_k evenly distributed between 0 and 2Hz:

$$u(t) = \sum_{k=1}^{d} a_k cos(2\pi\omega_k t + \phi_k) \tag{47}$$

where $a_k = 0.2$, $d = 50$ and the phase shift between each wave was defined according to the Schroeder choice (Schroeder, 1970):

$$\phi_k = \phi_1 - \frac{k(k-1)}{d}\pi \tag{48}$$

where $\phi_1 = 0$.

The continuous-time system was simulated using a 4^{th} order Runge-Kutta method and signals were sampled at varying frequencies: 10Hz, 20Hz, 40Hz, 80Hz, 160Hz and 320Hz. The simulation step-length was defined as 1/320. The range of sampling frequencies was chosen to illustrate the effect of relative slow and fast-sampling on structure selection and parameter estimation.

5.2 Expected discrete-time description of the Van-der-Pol oscillator

In the fast-sampling limit the non-linear δ-domain description is the equivalent of the continuous-time system. Hence at sampling frequencies below the continuous-time limit it may be expected that this equivalent δ-domain description dominates the structure selection. The structural equivalent of the continuous-time system in the δ-domain, in the fast-sampling limit, is

$$\delta^2 y(t) = \left[\zeta_1 - \zeta_2 y^2(t)\right] \delta y(t) - \zeta_3 y(t) + \zeta_4 u(t) \tag{49}$$

This model description can be mapped directly to the q-domain using the expression (26). Mapping (49) to the q-domain, after appropriate backward shifting, leads to the description

$$\begin{aligned} y(t) =& \theta_1 y(t-1) - \theta_2 y(t-2) + \theta_3 y^2(t-2) y(t-1) \\ &+ \theta_4 y^3(t-2) + \theta_5 u(t-2) \end{aligned} \tag{50}$$

The model structure above may indicate the non-linear basis functions likely to dominate the structure selection in the q-domain.

5.3 Identification

The structure of the non-linear model was detected by searching model orders $\{n_a, n_b, k\} \in \{1, \ldots, 5\}$ (where k

Iteration	Terms	Err	Param.
1	$y(t)$	0.31	-0.9981
2	$u(t)$	0.60	0.9956
3	$\delta y(t)y^2(t)$	0.072	-0.1991
4	$\delta y(t)$	0.011	0.1864
5	$\delta y(t)y(t)$	0.0050	-0.0757

Table 2
Significant terms selected for the δ-domain non-linear model of the Van-der-Pol oscillator at a sampling frequency of 40Hz.

was an input delay) and then iteratively increasing the maximum order of non-linearity p up to $p = 5$. The number of terms selected in the FRO procedure was a superset of 19 terms, which was pruned using the Bayesian Information Criterion (BIC) (Haber and Unbehauen, 1990).

The input-output order of the δ-domain model was detected as $n_a = 2$, $n_b = 0$ and delay $k = 0$. The identification of the difference equation model was performed using the usual backward shift operator q^{-1}, where $uq^{-1} = u(t-1)$. The input-output order of the q-domain model was identified as $n_a = 2$, $n_b = 1$ and $k = 1$. In the case of both the δ- and q-domain models the maximum order of non-linearity was selected as $p = 3$.

The results of the model identification procedure are shown for a sampling frequency of 40Hz, for the q-domain model in table 1 and the δ-domain model in table 2. These tables reveal the following points:

(1) The δ-domain model terms that are the fast-sampling limit equivalents of the continuous-time model have dominated the structure selection and the corresponding parameter estimates of those terms are similar to the continuous-time values.
(2) The FRO algorithm has attributed an Err of 1.00 to the q-domain term $y(t-1)$, which implies that close to 100% of the data is described by that one term. This potential problem was initially implied by the analysis in section 4.2.
(3) The q-domain model contains significantly more terms than the δ-domain model, suggesting that the use of the δ-operator has lead to a more parsimonious structure.

5.3.1 Parameter Estimation

Figure 1 demonstrates that the information matrix that was inverted during least squares parameter estimation was ill-conditioned for the q-domain model, even at low sampling frequencies. This numerical problem was exacerbated at higher sampling frequencies. In contrast the condition number of the δ-domain model was preserved across all sampling frequencies tested. This improvement in conditioning was notably present even at the slowest

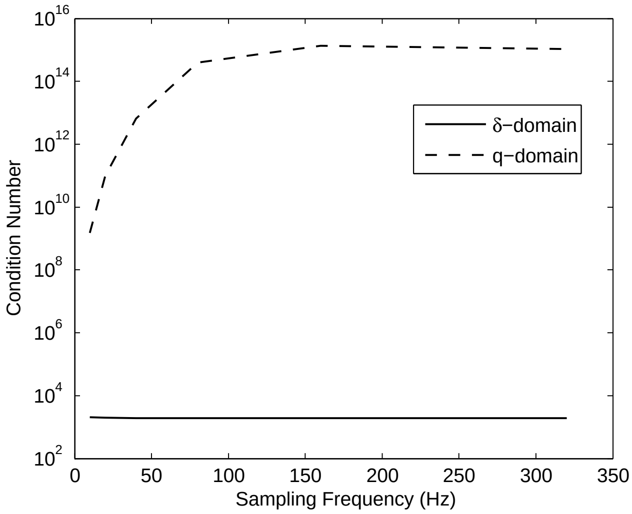

Fig. 1. The condition number of the information matrix that is inverted during least squares parameter estimation.

sampling frequency which was 10Hz (5 times the excitation frequency bandwidth).

The parameters of the δ-domain model terms contained in (49) were found to converge towards the continuous-time values as the sampling rate was increased, as shown in figure 2. This suggests that the δ-domain parameters converge towards the continuous-time values, even whilst the relevant terms are contained as a subset of the full discrete-time model. The parameter estimates of the q-domain model terms contained in (50) were found to have some similarity to the expected values, as shown in figure 2, but not the same properties of convergence as the δ-domain model. This result shows that the linear model δ-operator property of parameter convergence can extend, in some cases, to that of non-linear system identification.

5.3.2 Structure Detection

The model terms expected to dominate the structure selection were the δ equivalents of the continuous-time model and the corresponding q-domain model terms. It was found that the four expected δ-terms consistently dominated structure selection at all but the lowest sampling frequency (see table 3). In contrast the expected linear terms of the q-domain model dominated the structure detection, but the expected non-linear terms appeared at inconsistent intervals (see table 4).

Figure 3(a) demonstrates that the error variance of the model predicted output of the δ-domain model effectively converges after 5 model terms. This structure incorporates one additional term to the continuous-time system. The implication is that the process of sampling necessitates the inclusion of auxiliary terms in the iden-

Iteration	Terms	Err	Param.
1	$y(t-1)$	1.00	2.0006
2	$y(t-2)$	5.19 $\times 10^{-6}$	-1.0006
3	$u(t-1)$	5.30 $\times 10^{-11}$	9.7390 $\times 10^{-6}$
4	$y(t-1)u^2(t-1)$	2.03 $\times 10^{-13}$	1.6993 $\times 10^{-6}$
5	$y(t-2)u^2(t-1)$	6.14 $\times 10^{-14}$	-1.7023 $\times 10^{-6}$
6	$y^3(t-2)$	1.70 $\times 10^{-14}$	-0.5093
7	$y^3(t-1)$	4.20 $\times 10^{-14}$	0.5080
8	$y^2(t-2)$	3.98 $\times 10^{-14}$	0.0023
9	$y(t-1)y(t-2)$	2.18 $\times 10^{-14}$	-0.0044

Table 1
Significant terms selected for the q-domain non-linear model of the Van-der-Pol oscillator at a sampling frequency of 40Hz.

Freq. (Hz)	**Model Terms**				
	$y(t-1)$	$y(t-2)$	$u(t-2)$	$y^3(t-2)$	$y^2(t-2)y(t-1)$
10	1	2	3	5	7
20	1	2	3	5	7
40	1	2	3	6	10
80	1	2	3	6	10
160	1	2	3	9	10
320	1	2	3	9	10

Table 4
Iteration at which the relevant q-domain model term is selected in the FRO algorithm; with varying sample frequency.

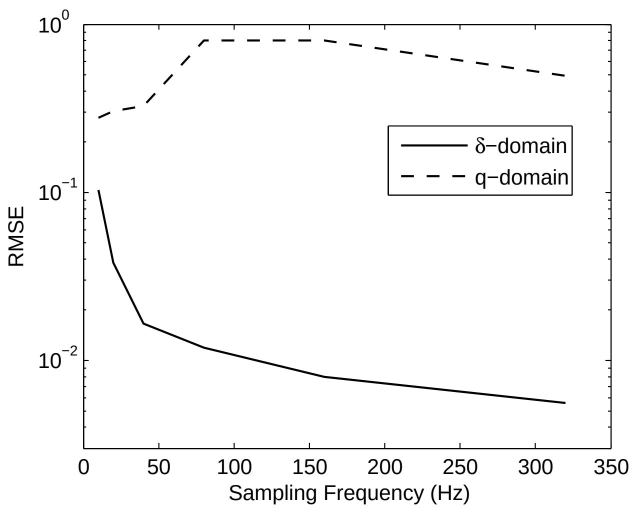

Fig. 2. Difference between continuous-time model parameters and the estimated parameters corresponding to the model terms expected to be obtained in the fast-sampling limit for both the q- and δ-domain models.

tified non-linear model. This does not appear to be affected by increasing the sampling rate. This is a contrasting result to that of parameter estimation, where the parameters do converge towards the continuous-time values as the sample rate increases.

Freq. (Hz)	**Model Terms**			
	$y(t)$	$u(t)$	$y^2(t)\delta y(t)$	$\delta y(t)$
10	1	2	3	6
20	1	2	3	4
40	1	2	3	4
80	1	2	3	4
160	1	2	3	4
320	1	2	3	4

Table 3
Iteration at which the relevant δ-domain model term is selected in the FRO algorithm; with varying sample frequency.

Figure 3(b) shows how the prediction error variance of the q-domain model has a tendency to increase at higher frequencies for a given number of model terms. This is in contrast to the δ-domain model; figure 3(a) shows that the error variance is decreased at higher sampling frequencies. This implies that the use of the δ-operator in non-linear system identification leads to more consistent performance across a range of sampling frequencies than the q-operator.

The structure of the non-linear model in both domains was detected by the use of the Err metric. In general it would be expected that the significant model terms

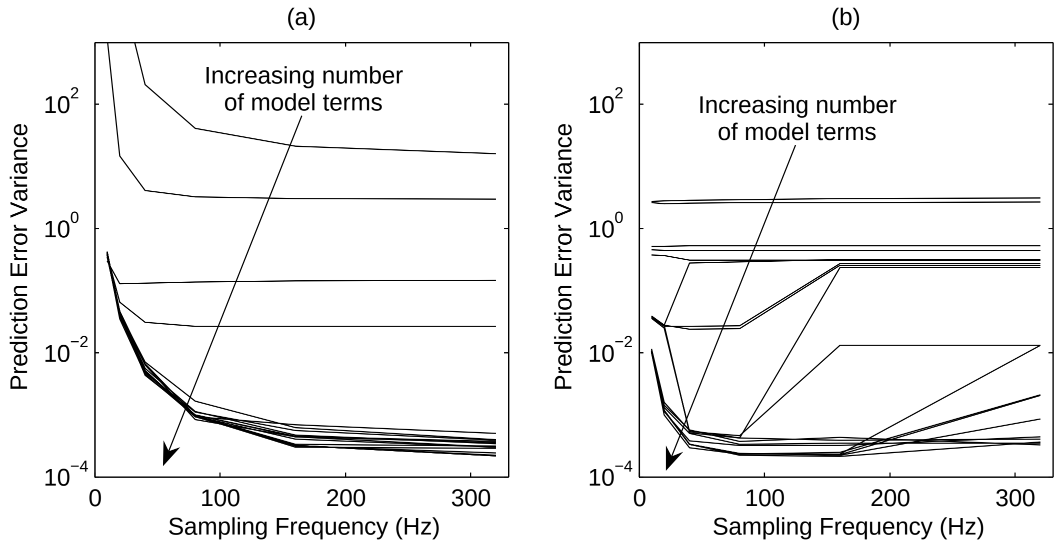

Fig. 3. Prediction error variance at increasing sampling frequencies, where each line corresponds to a model with a given number of terms from 1 to 19 in the order selected by the FRO algorithm: (a) δ-domain and (b) q-domain.

would have high Err and vice-versa. Therefore the sum of error reduction ratios (SERR) should be an indicator of model structure. Figure 4(a) indicates that there is a progressive usefulness in the selection of the first four terms in the δ-domain model, each of which have a direct correspondence to the continuous-time model (except in the case of the lowest sampling frequency). This progressive usefulness of model terms is not indicated in figure 4(b), for the q-domain model, where the SERR converges towards unity after the first term is selected. Note that this was to be expected from the analysis performed in section 4.2.

The BIC was used to truncate the ordered selection of model terms detected by the FRO algorithm. Figure 4(c) shows that the use of the δ-operator lead to the selection of fewer terms at all sampling frequencies compared to the q-domain model, where the corresponding results are shown in figure 4(d). The use of the δ-operator in the test scenario presented here has consistently lead to a more parsimonious structure at fast and slow sampling frequencies; this implies that in general the use of the δ-operator may lead to a more parsimonious description for certain model classes in non-linear system identification problems.

6 Conclusions

This investigation has shown the correspondence between polynomial non-linear models described in the q- and δ-domains. This exact relationship implies the potential theoretical equivalence of the use of either operator in non-linear system identification. However, the practical application of non-linear modelling techniques (e.g. FRO structure selection) has highlighted the fact that differences in identification (between q and δ) may arise, presumably resulting from numerical issues. Improvements from using the δ-operator in the identification of a deterministic continuous-time non-linear system have been demonstrated, focusing on: (i) the convergence of the parameter estimates of the non-linear δ-domain model to the continuous-time values (ii) the consistency of the detection procedure in terms of structural linking to the continuous-time system and consistent selection of auxiliary model terms that contribute to an accurate system description at higher sampling frequencies, and (iii) the parsimonious model description.

7 Acknowledgements

The authors would like to acknowledge the anonymous reviewers and the Editor for System Parameter Estimation for their valuable comments.

References

Billings, S. A., Aguirre, L. A., 1995. Effects of the sampling time on the dynamics and identification of non-linear models. International Journal of Bifurcation and Chaos 5, 1541–1556.

Chadwick, M. A., Kadirkamanathan, V., Billings, S. A., 2006. Analysis of fast-sampled non-linear systems: Generalised frequency response functions for δ-operator models. Signal Processing 86, 3246–3257.

Chen, S., Billings, S. A., Luo, W., 1989. Orthogonal least squares methods and their application to non-linear system identification. International Journal of Control 50, 1873–1896.

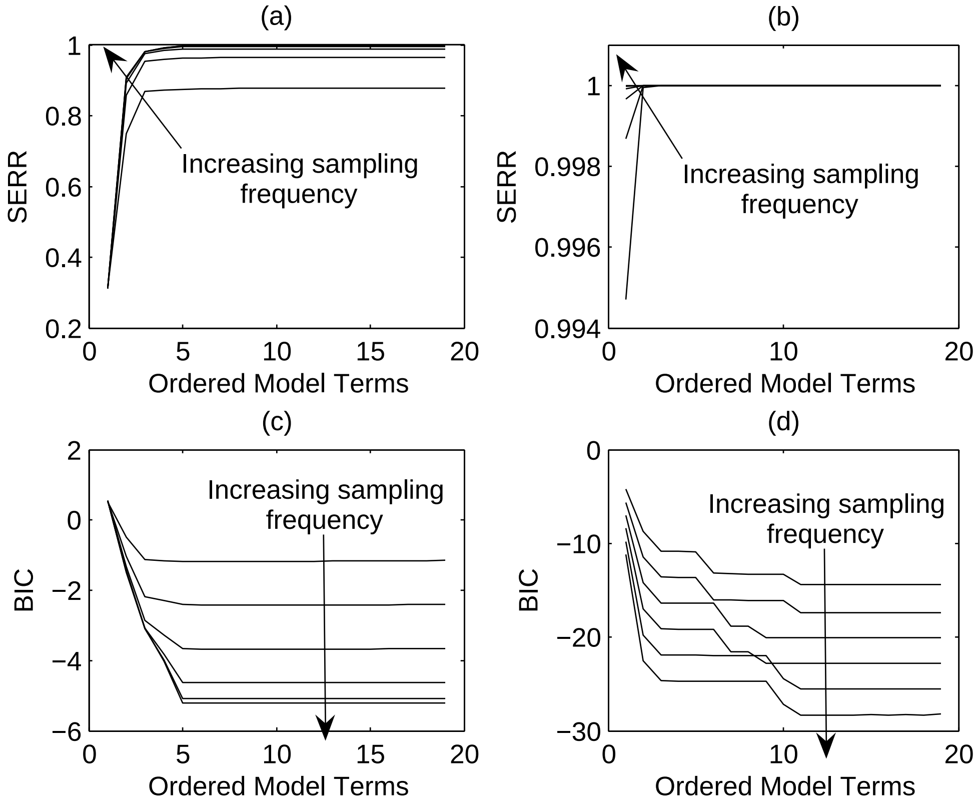

Fig. 4. (a) Sum of error reduction ratios of the δ-domain model. (b) Sum of error reduction ratios of the q-domain model. (c) Bayesian information criterion of model terms in the δ-domain model. (d) Bayesian information cost of model terms in the q-domain model.

Fan, H., De, P., 2001. High speed adaptive signal processing using the delta operator. Digital Signal Processing 11, 3–34.

Fan, H., Soderstrom, T., Mossberg, M., Carlsson, B., Yuanjie, Z., 1999. Estimation of continuous-time AR process parameters from discrete-time data. IEEE Transactions on Signal Processing 47, 1232–1244.

Goodwin, G. C., Middleton, R. H., Poor, H. V., 1992. High-speed digital signal processing and control. Proceedings of the IEEE 80, 240–259.

Haber, R., Unbehauen, H., 1990. Structure identification of nonlinear dynamic systems - a survey on input/output approaches. Automatica 26, 651–677.

Kuznetsov, A., Bowyer, R., Clarke, D. W., 1999. Estimation of multiple order models in the δ-domain. International Journal of Control 72, 629–642.

Larsson, E. K., Mossberg, M., Soderstrom, T., 2006. An overview of important practical aspects of continuous-time ARMA system identification. Circuits, Systems and Signal Processing 25 (1), 17–46.

Lauritsen, M. B., Rostgaard, M., Poulsen, N. K., 1997. GPC using delta-domain emulator-based approach. International Journal of Control 68, 219–232.

Leontaritis, I., Billings, S. A., 1985. Input-output parametric models for non-linear systems Part 1: deterministic nonlinear systems. International Journal of Control 2, 303–328.

Li, Q., Fan, H., 1997. On the properties of information matrices of delta-operator based adaptive signal processing algorithms. IEEE Trans. Signal Processing 45, 2454–2467.

Liu, G. P., Billings, S. A., Kadirkamanathan, V., 2000. Nonlinear system identification using wavelet networks. International Journal of Systems Science 31, 1531 – 1541.

Middleton, R. H., Goodwin, G. C., 1986. Improved finite word length characteristics in digital control using the delta operator. IEEE Trans. Automatic Control 31, 1015–1021.

Middleton, R. H., Goodwin, G. C., 1990. Digital Control and Estimation: A Unified Approach. Prentice Hall, Englewood Cliffs, NJ.

Monaco, S., Normand-Cyrot, D., 1995. A unified representation for nonlinear discrete-time and sampled dynamics. Journal of Mathematical Systems, Estimation and Control 5, 1–27.

Neuman, C., 1993. Transformations between delta and forward shift operator transfer function models. IEEE Trans. Systems, Man and Cybernetics 23, 295–296.

Schroeder, M. R., 1970. Synthesis of low-peak-factor signals and binary sequences with low autocorrelation. IEEE Transactions on Information Theory 16.

Shim, K., Sawan, M. E., 2006. Singularly perturbed unified time systems with low sensitivity to model reduction using delta operators. International Journal of Systems Science 37 (4), 243–251.

Soderstrom, T., Fan, H., Carlsson, B., Bigi, S., 1997. Least squares parameter estimation of continuous-time ARX models from discrete-time data. IEEE Transactions on Automatic Control 42, 659–673.

Suchomski, P., 2003. J-lossless and extended J-lossless factorization approach for delta-domain H-infinity control. International Journal of Control 76, 794–809.

Yuz, J. I., Goodwin, G. C., 2005. On sampled-data models for nonlinear systems. IEEE Transactions on Automatic Control 50.

A Proof of Lemma 3.1

Consider a single polynomial NARX model term $\chi(t)$ with non-linear order p, mapped to the δ-domain by substitution of $q = 1 + T\delta$ in (3),

$$\prod_{j=1}^{p} q^{n_j}\phi_j(t) = \prod_{j=1}^{p} (1+T\delta)^{n_j}\,\phi_j(t) \tag{A.1}$$

The terms arising from the RHS of this mapping can be be expressed as the ordered set $\mathcal{M}_{p,n_p}$,

$$\mathcal{M}_{p,n_p} =$$
$$\left\{\begin{array}{cccc}
c_{0,1}\delta^0\phi_1(t) & \times\; c_{0,2}\delta^0\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
c_{1,1}\delta^1\phi_1(t) & \times\; c_{0,2}\delta^0\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
c_{n_1,1}\delta^{n_1}\phi_1(t) & \times\; c_{0,2}\delta^0\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
c_{0,1}\delta^0\phi_1(t) & \times\; c_{1,2}\delta^1\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
c_{n_1,1}\delta^{n_1}\phi_1(t) & \times\; c_{1,2}\delta^1\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
c_{0,1}\delta^0\phi_1(t) & \times\; c_{n_2,2}\delta^{n_2}\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
c_{n_1,1}\delta^{n_1}\phi_1(t) & \times\; c_{n_2,2}\delta^{n_2}\phi_2(t) & \times\ldots\times & c_{0,p}\delta^0\phi_p(t)\\
c_{0,1}\delta^0\phi_1(t) & \times\; c_{0,2}\delta^0\phi_2(t) & \times\ldots\times & c_{1,p}\delta^1\phi_p(t)\\
c_{1,1}\delta^1\phi_1(t) & \times\; c_{0,2}\delta^0\phi_2(t) & \times\ldots\times & c_{1,p}\delta^1\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
c_{n_1,1}\delta^{n_1}\phi_1(t) & \times\; c_{0,2}\delta^0\phi_2(t) & \times\ldots\times & c_{1,p}\delta^1\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
\vdots & \vdots & \vdots & \vdots\\
c_{0,1}\delta^0\phi_1(t) & \times\; c_{n_2,2}\delta^{n_2}\phi_2(t) & \times\ldots\times & c_{n_p,p}\delta^{n_p}\phi_p(t)\\
\vdots & \vdots & \vdots & \vdots\\
c_{n_1,1}\delta^{n_1}\phi_1(t) & \times\; c_{n_2,2}\delta^{n_2}\phi_2(t) & \times\ldots\times & c_{n_p,p}\delta^{n_p}\phi_p(t)
\end{array}\right\} \tag{A.2}$$

The powers $m_{k,j}$, of $\delta^{m_{k,j}}$ from $\mathcal{M}_{p,n_p}$ can be defined for successive cross-product terms as (where the notation $m_{k,j}$ indicates the power of the k^{th} element and j^{th} cross-product term),

$$\begin{array}{rcl}
m_{k,1} & = & \bmod\ (k-1, n_1+1)\\
m_{k,2} & = & \bmod\ \left(\left\lfloor \frac{k-1}{n_1+1}\right\rfloor, n_2+1\right)\\
m_{k,3} & = & \bmod\ \left(\left\lfloor \frac{k-1}{(n_1+1)(n_2+1)}\right\rfloor, n_3+1\right)\\
\vdots & & \vdots\\
m_{k,p} & = & \bmod\ \left(\left\lfloor \frac{k-1}{\prod_{i=1}^{p-1} n_i+1}\right\rfloor, n_p+1\right)
\end{array} \tag{A.3}$$

Hence an arbitrary element $\mathbf{m}_k(t)$ in the set $\mathcal{M}_{p,n_p}$ can be defined as

$$\mathbf{m}_k(t) = \prod_{j=1}^{p} c_{m_{k,j},j}\delta^{m_{k,j}}\phi_j(t) \tag{A.4}$$

The full mapping of a q-domain term to the δ-domain involves the summation of the rows of the set $\mathcal{M}_{p,n_p}$ after each row is scaled by T raised to the appropriate power. The appropriate power of T corresponds to the sum of the powers of δ in the row k, $\bar{m}_k = \sum_{j=1}^{p} m_{k,j}$. Hence the mapping of a single q-domain NARX term $\chi(t)$ is

$$\prod_{j=1}^{p} q^{n_j}\phi_j(t) = \sum_{k=1}^{r} T^{\bar{m}_k}\mathbf{m}_k(t) \tag{A.5}$$

□

B Proof of Lemma 3.2

Consider the expression of a single δ-domain polynomial NARX model term $\psi(t)$ of order p mapped to the q-domain by substitution of (7) in (3)

$$\prod_{j=1}^{p} \delta^{n_j}\phi_j(t) = \prod_{j=1}^{p} \left(\frac{q-1}{T}\right)^{n_j}\phi_j(t) \tag{B.1}$$

The terms arising from the expansion of (B.1) can be written as an ordered set $\mathcal{M}^{-}_{p,n_p}$ similarly to $\mathcal{M}_{p,n_p}$, replacing the δ-operator by the q-operator and the binomial coefficients $c_{m_{k,j},j}$ by $c^{-}_{m_{k,j},j}$ and where each row of $\mathcal{M}^{-}_{p,n_p}$ is scaled by $T^{-\bar{m}_{l,k}}$. Hence the mapping of a single δ-domain NARX term $\psi(t)$ to the q-domain is

$$\prod_{j=1}^{p} \delta^{n_j}\phi_j(t) = \sum_{k=1}^{r} T^{\bar{m}_k}\mathbf{m}^{-}_k(t) \tag{B.2}$$

where $\mathbf{m}^{-}_k(t)$ is the k^{th} element in the set $\mathcal{M}^{-}_{p,n_p}$,

$$\mathbf{m}^{-}_k(t) = \prod_{j=1}^{p} c^{-}_{m_{k,j},j} q^{m_{k,j}}\phi_j(t). \tag{B.3}$$

□

C Proof of Theorem 3.1

The mapping of the l^{th} δ-domain to q-domain polynomial NARX model term follows from the extension of Lemma 3.1,

$$\prod_{j=1}^{p_l} q^{n_{l,j}}\phi_{l,j}(t) = \sum_{k=1}^{r_l} T^{\bar{m}_{l,k}} \prod_{j=1}^{p_l} c_{l,m_{l,k,j},j}\delta^{m_{l,k,j}}\phi_{l,j}(t). \tag{C.1}$$

where the application of Lemma 3.1 to the l^{th} model term leads to the definitions of (22), (23), (24) and (25). Hence the mapping of a full δ-domain model to the q-domain as defined in (21) follows immediately from the substitution of (C.1) in (4).

The preservation of the order of non-linearity when mapping a single term of order p from the q- to δ-domain follows from the definition of the set $\mathcal{M}_{p,n_p}$, where the number of cross-product terms in each element, resulting from the binomial expansion of (A.1), is exactly p.

The set of term orders $m_{k,j}$ contained within the set $\mathcal{M}_{p,n_p}$ is $\mathcal{N} = \{0,\ldots,n_1,\ldots,0,\ldots,n_p\}$. Correspondingly the set of term orders contained within the q-domain model term $\prod_{j=1}^{p} q^{n_j}\phi_j(t)$ is $\mathcal{N}_q = \{n_1,\ldots,n_p\}$, i.e $n_j \in \mathcal{N}_q$, for $j = 1,\ldots,p$. Clearly,

$$\max_{k=1,\ldots,r}\left(\mathcal{N}(k)\right) = \max_{j=1,\ldots,p}\left(\mathcal{N}_q(j)\right) \tag{C.2}$$

where the notation for a set $\mathcal{N}(k)$ indicates the k^{th} element of the set. By extension of (C.2) to the full model composed of w terms

$$\max_{l=1,\ldots,w}\left(\max_{k=1,\ldots,r_l}\left(\mathcal{N}^{(l)}(k)\right)\right) = \max_{l=1,\ldots,w}\left(\max_{j=1,\ldots,p_l}\left(\mathcal{N}^{(l)}_q(j)\right)\right) \tag{C.3}$$

which proves that the maximum time order of the model terms is preserved when mapping between domains. □

D Proof of Theorem 3.2

The mapping of the l^{th} q-domain to δ-domain polynomial NARX model term follows from the extension of Lemma 3.2,

$$\prod_{j=1}^{p_l} \delta^{n_{l,j}}\phi_{l,j}(t) = \sum_{k=1}^{r_l} T^{-\bar{m}_{l,k}} \prod_{j=1}^{p_l} c^{-}_{l,m_{l,k,j},j} q^{m_{l,k,j}}\phi_{l,j}(t). \tag{D.1}$$

where the application of Lemma 3.2 to the l^{th} model term leads to the definition of (27). Hence the mapping of a full q-domain model to the δ-domain as defined in (26) follows immediately from the substitution of (D.1) in (10).

The preservation of the order of non-linearity when mapping a single term of order p from the q- to δ-domain follows from the definition of the set $\mathcal{M}^{-}_{p,n_p}$, where the number of cross-product terms in each element, resulting from the binomial expansion of (B.1), is exactly p.

The set of term orders $m_{k,j}$ contained within the set $\mathcal{M}^{-}_{p,n_p}$ is $\mathcal{N}^{-} = \{0,\ldots,n_1,\ldots,0,\ldots,n_p\}$. Correspondingly the set of term orders contained within the δ-domain model term $\prod_{j=1}^{p} \delta^{n_j} \phi_j(t)$ is $\mathcal{N}_\delta = \{n_1, \ldots, n_p\}$, , i.e $n_j \in \mathcal{N}_\delta$, for $j = 1, \ldots, p$. Clearly,

$$\max_{k=1,\ldots,r} \left(\mathcal{N}^{-}(k)\right) = \max_{j=1,\ldots,p} \left(\mathcal{N}_\delta(j)\right) \tag{D.2}$$

and by extension to the full model composed of w terms

$$\max_{l=1,\ldots,w} \left(\max_{k=1,\ldots,r_l} \left(\mathcal{N}^{-(l)}(k)\right)\right) = \max_{l=1,\ldots,w} \left(\max_{j=1,\ldots,p_l} \left(\mathcal{N}_\delta^{(l)}(j)\right)\right) \tag{D.3}$$

which proves that the maximum time order of the model terms is preserved when mapping between domains.

□