Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

PARAMETER ESTIMATION BY QUASILINEARIZATION IN DIFFERENTIAL EQUATIONS WITH STATE-DEPENDENT DELAYS

Ferenc Hartung

Department of Mathematics
University of Pannonia
8201 Veszprém, P.O. Box 158, Hungary

Abstract. In this paper we study a parameter estimation method in functional differential equations with state-dependent delays using a quasilinearization technique. We define the method, prove its convergence under certain conditions, and test its applicability in numerical examples. We estimate infinite dimensional parameters such as coefficient functions, delay functions and initial functions in state-dependent delay equations. The method uses the derivative of the solution with respect to the parameters. The proof of the convergence is based on the Lipschitz continuity of the derivative with respect to the parameters.

1. **Introduction.** Estimation of unknown parameters in various classes of differential equations, and in particular in FDEs, has been investigated by many authors (see, e.g., [1, 2, 5, 6, 8, 20, 21, 23, 24, 27, 32]).

In this paper we consider the nonlinear scalar differential equation with state-dependent delay (SD-DDE)

$$\dot{x}(t) = f\Big(t, x_t, x(t-\tau(t,x_t,\xi)), \theta\Big), \qquad t \in [0,T] \tag{1.1}$$

with the associated initial condition

$$x(t) = \varphi(t), \qquad t \in [-r, 0]. \tag{1.2}$$

Throughout the manuscript $r > 0$ and $T > 0$ are fixed constants and $x_t\colon [-r,0] \to \mathbb{R}$, $x_t(s) := x(t+s)$ is the segment function. Let Θ and Ξ be normed linear spaces with norms $|\cdot|_\Theta$ and $|\cdot|_\Xi$, respectively, and suppose $\theta \in \Theta$ and $\xi \in \Xi$.

Here we consider the initial function φ, θ and ξ as parameters in the initial value problem (IVP) (1.1)-(1.2). In the next section we will define a parameter space Γ so that the IVP (1.1)-(1.2) has a unique solution $x(t,\gamma)$ corresponding to every $\gamma = (\varphi, \theta, \xi) \in \Gamma$ (see Theorem 2.1 below).

We assume that the parameter $\gamma = (\varphi, \xi, \theta) \in \Gamma$ is unknown, but there are measurements $X_0, X_1, \ldots, X_l$ of the solution at the points $t_0, t_1, \ldots, t_l \in [0,T]$. Our

---

2000 *Mathematics Subject Classification.* Primary: 34K05; Secondary: 93B30.

Key words and phrases. Parameter estimation, delay differential equation, state-dependent delay, quasilinearization, differentiability with respect to parameters.

This research was partially supported by the Hungarian National Foundation for Scientific Research Grant No. K101217.

goal is to find a parameter value which minimizes the least square cost function

$$J(\gamma) := \sum_{i=0}^{l} (x(t_i,\gamma) - X_i)^2 \tag{1.3}$$

over the parameter space Γ. Denote this infinite dimensional minimization problem by $\mathcal{P}$.

The quasilinearization (QL) method for solving problem $\mathcal{P}$ was introduced for ODEs in [3], and was applied to identify finite dimensional parameters in FDEs in [5] and [6]. The QL method was extended and numerically tested for SD-DDEs in [16]. The main goal of this paper is to prove the local convergence of the QL method for a class of SD-DDEs.

The QL method uses the derivative of the solution $x(t,\gamma)$ with respect to (wrt) γ, which is denoted by $D_2x(t,\gamma)$. The existence of this derivative is well-known under natural conditions for state-independent FDEs for a large class of parameters (see, e.g., [4, 13, 14, 29]). For SD-DDEs this is proved under restrictive assumptions (see [15, 17, 26, 27, 34, 35, 36, 37]); moreover, the differentiability was proved typically for the map $\gamma \mapsto x_t(\cdot,\gamma)$ using certain function norms on the state space. Recently the differentiability of the solutions $x(t,\gamma)$ wrt γ was proved in [19] under conditions which guarantee the existence of the derivative during the QL iteration.

The remaining part of this manuscript is organized as follows. In Section 2 we introduce our notations and hypotheses, discuss the well-posedness of the IVP (1.1)-(1.2), recall the results on the differentiability wrt parameters from [19], and show that under some conditions $D_2x(t,\gamma)$ is Lipschitz continuous in γ. In Section 3 we define the QL method, and in Section 4 we prove its local convergence. In Section 5 we show the applicability of the QL method for numerical examples in an SD-DDE.

2. **Well-posedness and differentiability wrt parameters.** A fixed norm on $\mathbb{R}^N$ and its induced matrix norm on $\mathbb{R}^{N\times N}$ are both denoted by $|\cdot|$. C denotes the Banach space of continuous functions $\psi\colon [-r,0]\to\mathbb{R}$ equipped with the norm $|\psi|_C = \max\{|\psi(\zeta)|\colon \zeta\in[-r,0]\}$. C^1 is the space of continuously differentiable functions $\psi\colon [-r,0]\to\mathbb{R}$ where the norm is defined by $|\psi|_{C^1} = \max\{|\psi|_C, |\dot\psi|_C\}$. L^∞ is the space of Lebesgue-measurable functions $\psi\colon [-r,0]\to\mathbb{R}$ which are essentially bounded. The norm on L^∞ is denoted by $|\psi|_{L^\infty} = \operatorname{ess\,sup}\{|\psi(\zeta)|\colon \zeta\in[-r,0]\}$. $W^{1,\infty}$ denotes the Banach space of absolutely continuous functions $\psi\colon [-r,0]\to\mathbb{R}$ of finite norm defined by

$$|\psi|_{W^{1,\infty}} := \max\left\{|\psi|_C, |\dot\psi|_{L^\infty}\right\}.$$

We note that $W^{1,\infty}$ is equal to the space of Lipschitz continuous functions from $[-r,0]$ to $\mathbb{R}$. The subset of $W^{1,\infty}$ consisting of those functions which have absolutely continuous first derivative and essentially bounded second derivative is denoted by $W^{2,\infty}$, where the norm is defined by

$$|\psi|_{W^{2,\infty}} := \max\left\{|\psi|_C,\ |\dot\psi|_C,\ |\ddot\psi|_{L^\infty}\right\}.$$

If the domain or the range of the functions is different from $[-r,0]$ and $\mathbb{R}$, respectively, we will use a more detailed notation. E.g., $C(X,Y)$ denotes the space of continuous functions mapping from X to Y. Finally, $\mathcal{L}(X,Y)$ denotes the space of bounded linear operators from X to Y, where X and Y are normed linear spaces.

An open ball in the normed linear space X centered at a point $x \in X$ with radius δ is denoted by $\mathcal{B}_X(x;\,\delta) := \{y \in Y\colon\ |x-y| < \delta\}$.

We introduce the parameter space

$$\Gamma := W^{1,\infty} \times \Theta \times \Xi$$

equipped with the product norm $|\gamma|_\Gamma := |\varphi|_{W^{1,\infty}} + |\theta|_\Theta + |\xi|_\Xi$ for $\gamma = (\varphi, \theta, \xi) \in \Gamma$.

For the well-posedness and differentiability results we assume

(A1) (i) $f\,:\,\mathbb{R} \times C \times \mathbb{R} \times \Theta \to \mathbb{R}$ is continuous;

(ii) $f(t, \psi, u, \theta)$ is Lipschitz continuous in ψ, u and θ, i.e., there exists a constant $L_1 \geq 0$ such that

$$|f(t, \psi, u, \theta) - f(t, \bar\psi, \bar u, \bar\theta)| \leq L_1\Big(|\psi - \bar\psi|_C + |u - \bar u| + |\theta - \bar\theta|_\Theta\Big),$$

for $t \in [0, T]$, $\psi, \bar\psi \in C$, $u, \bar u \in \mathbb{R}$ and $\theta, \bar\theta \in \Theta$;

(iii) $f\,:\,\mathbb{R} \times C \times \mathbb{R} \times \Theta \to \mathbb{R}$ is continuously differentiable wrt its second, third and fourth arguments;

(A2) (i) $\tau\,:\,\mathbb{R} \times C \times \Xi \to [0, r] \subset \mathbb{R}$ is continuous;

(ii) $\tau(t, \psi, \xi)$ is Lipschitz continuous in ψ and ξ, i.e., there exists a constant $L_2 \geq 0$ such that

$$|\tau(t, \psi, \xi) - \tau(t, \bar\psi, \bar\xi)| \leq L_2\Big(|\psi - \bar\psi|_C + |\xi - \bar\xi|_\Xi\Big)$$

for $t \in [0, T]$, $\psi, \bar\psi \in C$, $\xi, \bar\xi \in \Xi$;

(iii) $\tau\,:\,[0, T] \times C \times \Xi \to \mathbb{R}$ is continuously differentiable wrt its second and third arguments.

The well-posedness of several classes of SD-DDEs was studied in many papers, see, e.g., [10, 25, 26, 36, 37]. The next result is a variant of a result from [17] where the initial time was also considered as a parameter, but the parameters θ and ξ were not included in the equation. The proof is similar to that of Theorem 3.1 in [17] (see also the analogous proof of Theorem 3.2 of the neutral case in [18]), therefore it is omitted here. Note that in [17] and [18] local Lipschitz continuity was assumed on f and τ. In this manuscript global Lipschitz continuity is assumed for simplicity of the presentation.

Theorem 2.1. *Assume (A1) (i), (ii), (A2) (i), (ii), and let $\hat\gamma \in \Gamma$. Then there exist $\delta > 0$, $0 < \alpha \leq T$, N and L such that*

(i) *for all $\gamma = (\varphi, \theta, \xi) \in P := \mathcal{B}_\Gamma(\hat\gamma;\,\delta)$ the IVP (1.1)-(1.2) has a unique solution $x(t, \gamma)$ on $[-r, \alpha]$;*

(ii) *$x_t(\cdot, \gamma) \in W^{1,\infty}$ for $\gamma \in P$ and $t \in [0, \alpha]$, and*

$$|x_t(\cdot, \gamma)|_{W^{1,\infty}} \leq N, \qquad \gamma \in P,\ t \in [0, \alpha], \tag{2.1}$$

and

$$|x_t(\cdot, \gamma) - x_t(\cdot, \bar\gamma)|_{W^{1,\infty}} \leq L|\gamma - \bar\gamma|_\Gamma, \qquad \gamma \in P,\ t \in [0, \alpha]. \tag{2.2}$$

We assume that $\hat\gamma = (\hat\varphi, \hat\theta, \hat\xi) \in \Gamma$ is a fixed parameter, and its neighborhood P and the constant $\alpha > 0$ defined by Theorem 2.1 are also fixed thoroughout this paper.

Note that under the conditions of Theorem 2.1 the solutions, in general, are not C^1-functions, they are only $W^{1,\infty}$-functions. This lack of smoothness makes the study of differentiability wrt parameters technical.

We define the parameter set

$$P_1 := \{\gamma = (\varphi, \theta, \xi) \in P\colon\ x(\cdot, \gamma) \in X(\alpha, \xi)\}, \tag{2.3}$$

where P is defined in Theorem 2.1 and

$$X(\alpha, \xi) := \Big\{x \in W^{1,\infty}([-r, \alpha], \mathbb{R})\colon\ \operatorname{ess\,inf}\{\frac{d}{dt}(t - \tau(t, x_t, \xi))\colon\ \text{a.e. } t \in [0, \alpha^*]\} > 0\Big),$$

where $\alpha^* := \min\{r, \alpha\}$. Similar monotonicity condition of the time lag function was used in several papers in SD-DDEs ([7, 9, 19, 26, 31]), and, in general, it can be check if we know the solution. On the other hand, in some cases it can be guaranteed explicitly for large classes of parameters, see, e.g., [11].

We know (see [19] and [26]) that P_1 is an open subset of Γ, and it follows from the next theorem that for every $t \in [0, \alpha]$ and $\gamma \in P_1$ the derivative $D_2x(t, \gamma) \in \mathcal{L}(\Gamma, \mathbb{R})$ exists and is continuous.

Let $\gamma = (\varphi, \theta, \xi) \in P_1$ be fixed, and let $x(t) := x(t, \gamma)$. Then for a.e. $t \in [0, \alpha]$ we introduce the linear operator $L(t, x)\colon\ \Gamma \to \mathbb{R}$ by

$$\begin{aligned}
&L(t, x)(h^\varphi, h^\theta, h^\xi)\\
&:= \ D_2f(t, x_t, x(t - \tau(t, x_t, \xi)), \theta)h^\varphi + D_3f(t, x_t, x(t - \tau(t, x_t, \xi)), \theta)\\
&\qquad \times\Big[-\dot{x}(t - \tau(t, x_t, \xi))\Big(D_2\tau(t, x_t, \xi)h^\varphi + D_3\tau(t, x_t, \xi)h^\xi\Big) + h^\varphi(-\tau(t, x_t, \xi))\Big]\\
&\qquad + D_4f(t, x_t, x(t - \tau(t, x_t, \xi)), \theta)h^\theta, \qquad (h^\varphi, h^\theta, h^\xi) \in \Gamma.
\end{aligned} \tag{2.4}$$

It can be shown easily (see [19]) that $L(t, x)$ is a bounded linear operator for all t for which $\dot{x}(t - \tau(t, x_t, \xi))$ exists, i.e., for a.e. $t \in [0, \alpha]$.

For $\gamma \in P_1$ we define the variational equation associated to $x = x(\cdot, \gamma)$ as

$$\dot{z}(t) = L(t, x)(z_t, h^\theta, h^\xi) \qquad \text{a.e. } t \in [0, \alpha], \tag{2.5}$$

$$z(t) = h^\varphi(t), \qquad t \in [-r, 0], \tag{2.6}$$

where $h = (h^\varphi, h^\theta, h^\xi) \in \Gamma$ is fixed. The IVP (2.5)-(2.6) is a Carathéodory type linear delay equation. By its solution we mean a continuous function $z\colon\ [-r, \alpha] \to \mathbb{R}$ that is absolutely continuous on $[0, \alpha]$ and it satisfies (2.5) for a.e. $t \in [0, \alpha]$ and (2.6) for all $t \in [-r, 0]$. It is easy to show that the IVP (2.5)-(2.6) has a unique solution $z(t) = z(t, \gamma, h)$ for $t \in [-r, \alpha]$, $\gamma \in P_1$ and $h = (h^\varphi, h^\theta, h^\xi) \in \Gamma$, and that $z(t)$ is a bounded linear function of h for each fixed t and γ.

The next result shows continuous differentiability of the solution wrt the parameters. Note that this property was proved in [19] under a weaker assumption on the parameter set: instead of the montonicity of the time lag function ($\gamma \in P_1$) it was assumed a certain piecewise monotonicity property only.

Theorem 2.2 (see [19])**.** *Assume (A1) (i)–(iii), (A2) (i)–(iii), and let P_1 be defined by (2.3). Then the function*

$$\mathbb{R} \times \Gamma \supset [0, \alpha] \times P_1 \to \mathbb{R}, \qquad (t, \gamma) \mapsto x(t, \gamma)$$

is continuously differentiable wrt γ, and

$$D_2x(t, \gamma)h = z(t, \gamma, h), \qquad h \in \Gamma,\ t \in [0, \alpha],\ \gamma \in P_1, \tag{2.7}$$

where $z(t, \gamma, h)$ is the solution of the IVP (2.5)-(2.6) for $t \in [0, \alpha]$, $\gamma \in P_1$ and $h \in \Gamma$.

Moreover, there exists a constant $N_1 \geq 0$ such that

$$|D_2x(t, \gamma)h| \leq N_1|h|_\Gamma, \qquad h \in \Gamma,\ t \in [0, \alpha],\ \gamma \in P_1. \tag{2.8}$$

Proof. For the proof of (2.7) we refer to Theorem 4.7 and Remark 4.8 in [19]. For the proof of (2.8) see Lemma 4.3 in [19].

To show continuity of $D_2x(t,\gamma)$ wrt γ let $\gamma \in P_1$ be fixed, and let $h_k = (h_k^\varphi, h_k^\theta, h_k^\xi) \in \Gamma$ $(k \in \mathbb{N})$ be a sequence such that $|h_k|_\Gamma \to 0$ as $k \to \infty$ and $\gamma + h_k \in P_1$ for $k \in \mathbb{N}$. For a fixed $h = (h^\varphi, h^\theta, h^\xi) \in \Gamma$ we define the short notations $x^k(t) := x(t, \gamma + h_k)$, $x(t) := x(t,\gamma)$, $u^k(t) := t - \tau(t, x_t^k, \xi + h_k^\xi)$, $u(t) := t - \tau(t, x_t, \xi)$, $z^{k,h}(t) := z(t, \gamma + h_k, h)$ and $z^h(t) := z(t, \gamma, h)$. It was show in [19] that there exists a closed subset $M_1 \subset C$ which is also a bounded and convex subset of $W^{1,\infty}$ and $M_2 \subset \mathbb{R}$ closed and bounded interval such that $x_t, x_t^k \in M_1$ and $x(u(t)), x^k(u^k(t)) \in M_2$ for $t \in [0,\alpha]$ and $k \in \mathbb{N}$. Let $M_3 := \{\theta + \nu h_k^\theta : \nu \in [0,1],\ k \in \mathbb{N}\}$ and $M_4 := \{\xi + \nu h_k^\xi \colon \nu \in [0,1],\ k \in \mathbb{N}\}$. Then $M_1 \subset C$, $M_2 \subset \mathbb{R}$, $M_3 \subset \Theta$ and $M_4 \subset \Xi$ are compact subsets of the respective spaces.

It was shown in Lemma 4.5 of [19] that the functions $z^{k,h}$ and z^h satisfy

$$|z^{k,h}(t) - z^h(t)| \le c_{1,k} N_1 |h|_\Gamma, \qquad t \in [0,\alpha], \tag{2.9}$$

where

$$\begin{aligned}
c_{1,k} &:= \alpha(N_2+1)c_{0,k} + L_1L_2(N_2+1)\int_0^\alpha |\dot{x}(u^k(s)) - \dot{x}(u(s))|\, ds,\\
c_{0,k} &:= N_0\Omega_f\Big(K_3|h_k|_\Gamma\Big) + L_1L_2L|h_k|_\Gamma + L_1N\Omega_\tau\Big((L+1)|h_k|_\Gamma\Big) + L_1K_0|h_k|_\Gamma,\\
\Omega_f(\varepsilon) &:= \max_{i=2,3,4}\sup\Big\{|D_if(t,\psi,u,\theta) - D_if(t,\tilde\psi,\tilde u,\tilde\theta)|_{\mathcal{L}(Y_i,\mathbb{R})}\colon\\
&\qquad\qquad |\psi - \tilde\psi|_C + |u - \tilde u| + |\theta - \tilde\theta|_\Theta \le \varepsilon,\ \ t \in [0,\alpha],\ \psi, \tilde\psi \in M_1,\\
&\qquad\qquad u, \tilde u \in M_2,\ \theta, \tilde\theta \in M_3\Big\},\\
\Omega_\tau(\varepsilon) &:= \max_{i=2,3}\sup\Big\{|D_i\tau(t,\psi,\xi) - D_i\tau(t,\bar\psi,\bar\xi)|_{\mathcal{L}(Z_i,\mathbb{R})}\colon |\psi - \bar\psi|_C + |\xi - \bar\xi|_\Xi \le \varepsilon,\\
&\qquad\qquad t \in [0,\alpha],\ \psi, \bar\psi \in M_1,\ \xi, \bar\xi \in M_4\Big\},
\end{aligned}$$

where $Y_2 := C$, $Y_3 := \mathbb{R}$, $Y_4 := \Theta$, $Z_2 := C$ and $Z_3 := \Xi$, and $K_0 \ge 0$ and $N_2 \ge 0$ are certain costants. The continuity of the partial derivatives of f and τ yield that $\Omega_f(\varepsilon) \to 0$ and $\Omega_\tau(\varepsilon) \to 0$ as $\varepsilon \to 0+$, and hence $c_{0,k} \to 0$ as $k \to \infty$. It was argued in [19] that the integral in the definition of $c_{1,k}$ goes to 0 as $k \to \infty$, which implies the continuity of the map $\Gamma \supset P_1 \ni \gamma \mapsto D_2x(t,\gamma) \in \mathcal{L}(\Gamma,\mathbb{R})$. The continuity of $D_2x(t,\gamma)$ wrt to t can also be shown. For the details see [19]. □

Lemma 2.6 below shows that, under additional conditions, the function $\Gamma \supset P_1 \ni \gamma \mapsto D_2x(t,\gamma) \in \mathcal{L}(\Gamma,\mathbb{R})$ is Lipschitz continuous. We will need the following additional assumptions.

(A1) (iv) we take L_1 in (A1) (ii) to also be a Lipschitz constant of f with respect to t, i.e.,

$$|f(t,\psi,u,\theta) - f(\bar t,\psi,u,\theta)| \le L_1|t - \bar t|$$

for $t, \bar t \in [0,T]$, $\psi \in C$, $u \in \mathbb{R}$ and $\theta \in \Theta$;

(v) D_2f, D_3f and D_4f are Lipschitz continuous wrt all of their arguments, i.e., there exists $L_3 \ge 0$ such that

$$|D_if(t,\psi,u,\theta) - D_if(\bar t,\bar\psi,\bar u,\bar\theta)|_{\mathcal{L}(Y_i,\mathbb{R})} \le L_3\Big(|t - \bar t| + |\psi - \bar\psi|_C + |u - \bar u| + |\theta - \bar\theta|_\Theta\Big)$$

for $i = 2, 3, 4$, $t, \bar{t} \in [0, T]$, $\psi, \bar{\psi} \in C$, $u, \bar{u} \in \mathbb{R}$ and $\theta, \bar{\theta} \in \Theta$, where $Y_2 := C$, $Y_3 := \mathbb{R}$ and $Y_4 := \Theta$;

(A2) (iv) we take L_2 in (A2) (ii) to also be a Lipschitz constant of τ with respect to t, i.e.,

$$|\tau(t, \psi, \xi) - \tau(\bar{t}, \psi, \xi)| \leq L_2 |t - \bar{t}|$$

for $t, \bar{t} \in [0, T]$, $\psi \in C$, $\xi \in \Xi$;

(v) there exists $L_4 \geq 0$ such that

$$\Big|\frac{d}{dt}\tau(t, y_t, \xi) - \frac{d}{dt}\tau(t, \bar{y}_t, \bar{\xi})\Big| \leq L_4\Big(|y_t - \bar{y}_t|_{W^{1,\infty}} + |\xi - \bar{\xi}|_{\Xi}\Big), \quad \text{a.e. } t \in [0, \alpha],$$

where $\xi, \bar{\xi} \in \Xi$, and $y, \bar{y} \in W^{1,\infty}([-r, \alpha], \mathbb{R})$;

(vi) $D_2\tau$ and $D_3\tau$ are Lipschitz continuous wrt all arguments, i.e., there exists a constant $L_5 \geq 0$ such that

$$|D_i\tau(t, \psi, \xi) - D_i\tau(\bar{t}, \bar{\psi}, \bar{\xi})|_{\mathcal{L}(Z_i, \mathbb{R})} \leq L_5\Big(|t - \bar{t}| + |\psi - \bar{\psi}|_C + |\xi - \bar{\xi}|_{\Xi}\Big)$$

for $i = 2, 3$, $t, \bar{t} \in [0, T]$, $\psi, \bar{\psi} \in C$, $\xi, \bar{\xi} \in \Xi$, where $Z_2 := C$ and $Z_3 := \Xi$.

First we recall the following technical result from [19].

Lemma 2.3. *Assume (A1) (i), (ii), (A2) (i),(ii), $\gamma = (\varphi, \xi, \theta) \in P$, and $h_k = (h_k^\varphi, h_k^\xi, h_k^\theta) \in \Gamma$ is a sequence such that $\gamma + h_k \in P$ for $k \in \mathbb{N}$ and $|h_k|_\Gamma \to 0$ as $k \to \infty$. Let $x(t) := x(t, \gamma)$, $x^k(t) := x(t, \gamma + h_k)$ be the corresponding solutions of the IVP (1.1)-(1.2), and $u^k(s) := t - \tau(t, x_t^k, \xi + h_k^\xi)$ and $u(t) := t - \tau(t, x_t, \xi)$. Then there exists $K_0 \geq 0$ such that*

$$|u^k(t) - u(t)| \leq K_0 |h_k|_\Gamma, \qquad t \in [0, \alpha], \quad k \in \mathbb{N}. \tag{2.10}$$

If, in addition, (A2) (iv) holds, then $u, u^k \in W^{1,\infty}([0, \alpha], \mathbb{R})$, and if (A2) (v) is also satisfied, then there exist $K_1 \geq 0$ and $K_2 \geq 0$ such that

$$|u^k - u|_{W^{1,\infty}([0,\alpha],\mathbb{R})} \leq K_1 |h_k|_\Gamma, \qquad k \in \mathbb{N} \tag{2.11}$$

and

$$|x^k(u^k(t)) - x(u(t))| \leq K_2 |h_k|_\Gamma, \qquad k \in \mathbb{N}, \quad t \in [0, \alpha]. \tag{2.12}$$

Later we will need the following estimate, which is an easy consequence of assumption (A2) (ii) and (2.1):

$$\begin{aligned} |x(u(t)) - x(u(\bar{t}))| &\leq N|u(t) - u(\bar{t})| \\ &\leq NL_2(|t - \bar{t}| + |x_t - x_{\bar{t}}|_C) \\ &\leq NL_2(1 + N)|t - \bar{t}|, \qquad t, \bar{t} \in [0, T]. \end{aligned} \tag{2.13}$$

Lemma 2.4. *Assume (A1) (i)–(iv), (A2) (i)–(iv) and $\gamma = (\varphi, \theta, \xi) \in P$ is such that $\varphi \in W^{2,\infty}$. Then there exists $K_4 = K_4(\gamma) \geq 0$ such that the solution $x(t) = x(t, \gamma)$ of the IVP (1.1)-(1.2) satisfies*

$$|\dot{x}(t) - \dot{x}(\bar{t})| \leq K_4 |t - \bar{t}| \qquad \text{for } t, \bar{t} \in [-r, 0) \quad \text{and} \quad t, \bar{t} \in (0, \alpha]. \tag{2.14}$$

Proof. The Mean Value Theorem and the definition of the $W^{2,\infty}$-norm yield

$$|\dot x(t)-\dot x(\bar t)| = |\dot\varphi(t)-\dot\varphi(\bar t)| \le |\varphi|_{W^{2,\infty}}|t-\bar t|, \qquad t,\bar t\in[-r,0).$$

For $t,\bar t\in(0,\alpha]$ it follows from (A1) (ii), (iv), (A2) (ii), (iv), (2.1) and (2.13)

$$\begin{aligned}|\dot x(t)-\dot x(\bar t)| &= |f(t,x_t,x(u(t)),\theta)-f(\bar t,x_{\bar t},x(u(\bar t)),\theta)|\\ &\le L_1\Big(|t-\bar t|+|x_t-x_{\bar t}|_C+|x(u(t))-x(u(\bar t))|\Big)\\ &\le L_1\Big(1+N+NL_2(1+N)\Big)|t-\bar t|.\end{aligned}$$

Hence (2.14) is satisfied with $K_4:=\max\{|\varphi|_{W^{2,\infty}}, L_1[1+N+NL_2(1+N)]\}$. □

We will need the following class of initial functions in the next lemma.

Definition 2.5. *Let $PW^{2,\infty}$ denote the set of functions $\varphi\in W^{1,\infty}$ which are piecewise $W^{2,\infty}$-functions, i.e., there exists a finite mesh $-r=t_0<t_1<\ldots<t_{\ell+1}=0$ such that*

(i) *$\dot\varphi$ is Lipschitz continuous on the intervals (t_i,t_{i+1}) for $i=0,\ldots,\ell$, and*
(ii) *$\dot\varphi$ has continuous one-sided derivatives at t_i for $i=0,\ldots,\ell+1$.*

We define a norm on $PW^{2,\infty}$ by $|\varphi|_{PW^{2,\infty}}:=\max\{|\varphi|_C,|\dot\varphi|_{L^\infty},|\ddot\varphi|_{L^\infty}\}$.

Note that any function $\varphi\in PW^{2,\infty}$ is almost everywhere differentiable and twice differentiable, but $\dot\varphi$ may have discontinuity at the mesh points $t_1,\ldots,t_\ell$. A typical example of a $PW^{2,\infty}$-function is a spline function defined on $[-r,0]$.

The next lemma gives sufficient conditions under which $D_2x(t,\gamma)$ depends Lipschitz continuously on γ. This result will be essential to prove the convergence of the QL sequence in Section 4.

Lemma 2.6. *Assume (A1) (i)–(v), (A2) (i)–(vi), and $\gamma^*=(\varphi^*,\theta^*,\xi^*)\in P_1$. Then there exists $\delta^*>0$ such that for every $m\in\mathbb N$ and $K\ge 0$ there exists a nonnegative constant $N_3=N_3(\gamma^*,\delta^*,m,K)$ such that for every $\gamma=(\varphi,\theta,\xi)\in\mathcal B_\Gamma(\gamma^*;\delta^*)$ satisfying $\varphi\in PW^{2,\infty}$ with $|\varphi|_{PW^{2,\infty}}\le K$, and the number of points of discontinuity of $\dot\varphi$ in $(-r,0)$ is less or equal to m, there exists $\delta>0$ such that for every sequence $h_k\in\Gamma$ with $|h_k|_\Gamma\le\delta$ for $k\in\mathbb N$ and all $h\in\Gamma$ the functions $z^{k,h}(t):=z(t,\gamma+h_k,h)$ and $z^h(t):=z(t,\gamma,h)$ satisfy*

$$|z^{k,h}(t)-z^h(t)|\le|z^{k,h}_t-z^h_t|_C\le N_3|h_k|_\Gamma|h|_\Gamma, \qquad t\in[0,\alpha],\quad h\in\Gamma. \tag{2.15}$$

Proof. Since P_1 is an open subset of P (see [26] and [17]), there exists a $\delta_0>0$ such that $\mathcal B_\Gamma(\gamma^*;\delta_0)\subset P_1$. For a fixed $\gamma\in\mathcal B_\Gamma(\gamma^*;\delta_0)$ we define $x(t):=x(t,\gamma)$, $x^*(t):=x(t,\gamma^*)$, $u(t):=t-\tau(t,x_t,\xi)$ and $u^*(t):=t-\tau(t,x^*_t,\xi^*)$. Introduce

$$M^*:=\min\Big\{\operatorname*{ess\,inf}_{s\in[0,\alpha^*]}\dot u^*(s),\,1\Big\}.$$

Then $\gamma^*\in P_1$ yields $M^*>0$, and u^* is strictly monotone increasing on $[0,\alpha^*]$. Let $0<M<M^*$ be fixed. It follows from Lemma 2.3 that there exists $0<\delta^*\le\delta_0$ such that if $\gamma\in\mathcal B_\Gamma(\gamma^*;\delta^*)$, then $\dot u(s)\ge M$ for a.e. $s\in[0,\alpha^*]$, and, in particular, u is also strictly monotone increasing on $[0,\alpha^*]$.

Fix $m\in\mathbb N$ and $\gamma=(\varphi,\theta,\xi)\in\mathcal B_\Gamma(\gamma^*;\delta^*)$ such that $\varphi\in PW^{2,\infty}$ and the points of discontinuity of $\dot\varphi$ in $(-r,0)$ is less or equal to m, and let K be such that $|\varphi|_{PW^{2,\infty}}\le K$. Let $\delta_1\ge 0$ be such that $\mathcal B_\Gamma(\gamma;\delta_1)\subset\mathcal B_\Gamma(\gamma^*;\delta^*)$, and let $h_k\in\Gamma$ $(k\in\mathbb N)$ be a sequence satisfying $|h_k|_\Gamma\le\delta_1$ for $k\in\mathbb N$. Let $x^k(t):=x(t,\gamma+h_k)$

and $u^k(t) := t - \tau(t, x^k_t, \xi + h^\xi_k)$. Let $-r < t_1 < \cdots < t_\ell < 0$ be the points of discontinuity of $\dot\varphi$ (from Definition 2.5), and define $t_0 := -r$ and $t_{\ell+1} := 0$. Then by the assumption on γ we have $\ell \le m$.

The proof of Lemma 2.4 yields that $K_4^* := \max\{K, L_1[1 + N + NL_2(1 + N)]\}$ satisfies

$$|\dot x(t) - \dot x(\bar t)| \le K_4^* |t - \bar t| \qquad \text{for } t, \bar t \in (t_i, t_{i+1}), \quad i = 0, \dots, \ell, \quad t, \bar t \in (0, \alpha). \tag{2.16}$$

Let $\varepsilon_0 := \min\{t_{i+1} - t_i : i = 0, \dots, \ell\}$. Let $\delta_2 := \min\Big\{\delta_1, \frac{M\varepsilon_0}{K_0}\Big\}$. Then if $|h_k|_\Gamma < \delta_2$ for all $k \in \mathbb{N}$, then by (2.10) we have

$$|u^k(s) - u(s)| \le K_0 |h_k|_\Gamma \le M\varepsilon_0 \le \varepsilon_0, \qquad k \in \mathbb{N}, \quad s \in [0, \alpha^*]. \tag{2.17}$$

Since $u(0) \le 0$, there exist $s_i \in [0, \alpha^*]$ and $j \in \{0, 1, \dots, \ell + 1\}$ such that $u(s_i) = t_i$ for $i = j, \dots, \ell + 1$. By the strict monotonicity of u we have $0 \le s_j < \cdots < s_{\ell+1} \le \alpha^*$. Similarly, let $s_{k,i}$ and j_k be such that $u^k(s_{k,i}) = t_i$ for $i = j_k, \dots, \ell + 1$, $k \in \mathbb{N}$. We again have $0 \le s_{k,j_k} < \cdots < s_{k,\ell+1} \le \alpha^*$.

Next we show that if $|h_k|_\Gamma < \delta_2$ for $k \in \mathbb{N}$, then

$$|s_{k,i} - s_i| \le \frac{K_0}{M} |h_k|_\Gamma \le \varepsilon_0, \qquad i = \max(j, j_k), \dots, \ell + 1, \quad k \in \mathbb{N}. \tag{2.18}$$

First consider the case when $s_{k,i} \ge s_i$ for some $i \in \{\max(j, j_k), \dots, \ell+1\}$ and $k \in \mathbb{N}$. The definitions of M, δ^*, δ_1, δ_2, s_i and $s_{k,i}$ and (2.17) imply

$$M(s_{k,i} - s_i) \le u(s_{k,i}) - u(s_i) = u(s_{k,i}) - u^k(s_{k,i}) \le K_0 |h_k|_\Gamma \le M\varepsilon_0, \qquad k \in \mathbb{N}$$

for all $i = \max(j, j_k), \dots, \ell + 1$. We have then $0 \le s_{k,i} - s_i \le \varepsilon_0$. In the opposite case when $s_{k,i} < s_i$ we get the same way that $0 \le s_i - s_{k,i} \le \frac{K_0}{M} |h_k|_\Gamma \le \varepsilon_0$, which yields (2.18).

We distinguish 3 cases. Case (1): If $j = 0$, then $s_j = 0$, moreover, $j_k = 0$ and $s_{k,j_k} = 0$ for $u^k(0) = -r$, and $j_k = 1$ and $s_{k,j_k} > 0$ for $u^k(0) > -r$. Case (2): If $s_j = 0$ and $j > 0$, then $u(0) = t_j$, moreover, $j_k = j + 1$ and $s_{k,j+1} > 0$ for $u^k(0) > u(0)$, and $j_k = j$ and $s_{k,j} \ge 0$ for $u^k(0) \le u(0)$. Case (3): $s_j > 0$ and $j > 0$. Then $t_{j-1} < u(0) < t_j$, and let $\Delta := \min(u(0) - t_{j-1}, t_j - u(0))$ and $\delta_3 := \min\{\delta_2, \frac{\Delta}{K_0}\}$. Then if $|h_k|_\Gamma < \delta_3$ for all $k \in \mathbb{N}$, then $|u^k(s) - u(s)| \le K_0 |h_k|_\Gamma < \Delta$ for $s \in [0, \alpha^*]$, and hence $j_k = j$, and $u^k(s), u(s) \in (t_{j-1}, t_j)$ for $0 \le s < \min(s_j, s_{k,j})$.

Now we consider Case (3) above. Suppose $|h_k|_\Gamma < \delta_3$ for all $k \in \mathbb{N}$. Define $a_{k,i} := \min(s_i, s_{k,i})$ and $b_{k,i} := \max(s_i, s_{k,i})$ for $i = j, \dots, \ell+1$. Then for $i = j, \dots, \ell$ and $k \in \mathbb{N}$ we have

$$b_{k,i} - a_{k,i} = |s_i - s_{k,i}| \le \frac{K_0}{M} |h_k|_\Gamma, \tag{2.19}$$

$b_{k,i} < a_{k,i+1}$, and $u(s), u^k(s) \in (t_i, t_{i+1})$ for $s \in (b_{k,i}, a_{k,i+1})$. For definiteness suppose $(a_{k,i}, b_{k,i}) = (s_i, s_{k,i})$ (the opposite case is similar). Then for $s \in (a_{k,i}, b_{k,i})$ we have $u(s) \in (t_i, t_{i+1})$ and $u^k(s) \in (t_{i-1}, t_i)$. Therefore (2.16) and (2.10) imply

$$\begin{aligned}
&|\dot x(u(s)) - \dot x(u^k(s))| \\
&\quad \le |\dot x(u(s)) - \dot x(t_i+)| + |\dot x(t_i+) - \dot x(t_i-)| + |\dot x(t_i-) - \dot x(u^k(s))| \\
&\quad \le K_4^*(u(s) - t_i) + |\dot x(t_i+) - \dot x(t_i-)| + K_4^*(t_i - u^k(s)) \\
&\quad \le K_4^* |u(s) - u^k(s)| + |\dot x(t_i+) - \dot x(t_i-)| \\
&\quad \le K_4^* K_0 |h_k|_\Gamma + |\dot x(t_i+) - \dot x(t_i-)|.
\end{aligned} \tag{2.20}$$

Then (A1) (ii), (2.2) and (2.12) give for $t \in [0, \alpha]$

$$\begin{aligned}
|\dot{x}(t)| &\le |f(t, x_t, x(u(t)), \theta) - f(t, x_t^*, x^*(u^*(t)), \theta^*)| + |f(t, x_t^*, x^*(u^*(t)), \theta^*)| \\
&\le L_1(|x_t - x_t^*|_C + |x(u(t)) - x^*(u^*(t))| + |\theta - \theta^*|_\Theta) \\
&\quad + \max_{t\in[0,\alpha]} |f(t, x_t^*, x^*(u^*(t)), \theta^*)| \\
&\le L_1(L + K_2 + 1)|\gamma - \gamma^*|_\Gamma + \max_{t\in[0,\alpha]} |f(t, x_t^*, x^*(u^*(t)), \theta^*)| \\
&\le \widehat{K},
\end{aligned}$$

where $\widehat{K} := L_1(L+K_2+1)\delta^* + \max_{t\in[0,\alpha]} |f(t, x_t^*, x^* u^*(t)), \theta^*)|$. Then, in particular, $|\dot{x}(0+)| \le \widehat{K}$ for all $\gamma \in \mathcal{B}_\Gamma(\gamma^*; \delta^*)$, and so (2.20) yields for all $i = j, \dots, \ell$ and $k \in \mathbb{N}$

$$|\dot{x}(u(s)) - \dot{x}(u^k(s))| \le K_4^* K_0 |h_k|_\Gamma + 2K^*, \qquad s \in (a_{k,i}, b_{k,i}), \tag{2.21}$$

where $K^* := \max\{K, \widehat{K}\}$. Note that it is easy to check that (2.21) holds for the case $(a_{k,i}, b_{k,i}) = (s_{k,i}, s_i)$, too.

Therefore by (2.10), (2.16), (2.19), (2.21) and $\ell \le m$ we have

$$\begin{aligned}
&\int_0^{\alpha^*} |\dot{x}(u(s)) - \dot{x}(u^k(s))|\, ds \\
&\quad = \int_0^{a_{k,j}} |\dot{x}(u(s)) - \dot{x}(u^k(s))|\, ds + \sum_{i=j}^{\ell} \int_{a_{k,i}}^{b_{k,i}} |\dot{x}(u(s)) - \dot{x}(u^k(s))|\, ds \\
&\qquad + \sum_{i=j}^{\ell} \int_{b_{k,i}}^{a_{k,i+1}} |\dot{x}(u(s)) - \dot{x}(u^k(s))|\, ds + \int_{b_{k,\ell+1}}^{\alpha^*} |\dot{x}(u(s)) - \dot{x}(u^k(s))|\, ds \\
&\quad \le a_{k,j} K_4^* K_0 |h_k|_\Gamma + \sum_{i=j}^{\ell} (b_{k,i} - a_{k,i}) K_4^* K_0 |h_k|_\Gamma + \sum_{i=j}^{\ell} (b_{k,i} - a_{k,i}) 2K^* \\
&\qquad + \sum_{i=j}^{\ell} (a_{k,i+1} - b_{k,i}) K_4^* K_0 |h_k|_\Gamma + (\alpha^* - b_{k,\ell+1}) K_4^* K_0 |h_k|_\Gamma \\
&\quad \le \Big(\alpha^* K_4^* K_0 + m \frac{K_0}{M} 2K^*\Big) |h_k|_\Gamma.
\end{aligned} \tag{2.22}$$

Inequality (2.22) can be obtained similarly for the Cases (1) and (2).

Assumptions (A1) (v) and (A2) (vi) imply that $\Omega_f(\varepsilon) \le L_3 \varepsilon$ and $\Omega_\tau(\varepsilon) \le L_5 \varepsilon$ for $\varepsilon \ge 0$. Therefore the definitions of $c_{0,k}$, $c_{1,k}$ and (2.22) yield the existence of an $L^* \ge 0$ such that $c_{1,k} \le L^* |h_k|_\Gamma$ for all h_k satisfying $|h_k|_\Gamma < \delta$ for some $\delta > 0$. Then (2.15) follows from (2.9) with $N_3 := L^* N_1$. □

3. **Formulation of the quasilinearization method.** Following [28], we briefly show the derivation of the QL method. Let $X_0, X_1, \dots, X_l$ be measurements of the solution corresponding to an unknown parameter at the points $t_0, t_1, \dots, t_l \in [0, T]$. Let Γ^N be an N-dimensional subspace of the parameter space Γ, and let $\gamma_k = (\varphi_k, \theta_k, \xi_k) \in \Gamma^N$ be fixed, and consider the corresponding solution of the IVP (1.1)-(1.2), $x(t, \gamma_k)$. For a fixed $i \in \{0, 1, \dots, \ell\}$ take first order Taylor-approximation of $x(t_i, \gamma)$ around the parameter γ_k:

$$x(t_i, \gamma) \approx x(t_i, \gamma_k) + D_2 x(t_i, \gamma_k)(\gamma - \gamma_k),$$

and consider the approximate cost function restricted to the subspace Γ^N defined by

$$J^{k,N}(\gamma) := \sum_{i=0}^{l}\Big(x(t_i,\gamma_k) + D_2x(t_i,\gamma_k)(\gamma-\gamma_k) - X_i\Big)^2, \qquad \gamma\in\Gamma^N.$$

We solve the minimization problem $\mathcal{P}^{k,N}$:

$$\min_{\gamma\in\Gamma^N} J^{k,N}(\gamma).$$

Fix a basis $\{\chi_1^N,\ldots,\chi_N^N\}$ for the finite dimensional subspace Γ^N, and for $\gamma_k,\gamma\in\Gamma^N$ let

$$\gamma_k := \sum_{j=1}^{N} c_j^k\chi_j^N \qquad \text{and} \qquad \gamma := \sum_{j=1}^{N} c_j\chi_j^N.$$

We introduce the vectors $\mathbf{c}^k := (c_1^k,\ldots,c_N^k)^T\in\mathbb{R}^N$ and $\mathbf{c} := (c_1,\ldots,c_N)^T\in\mathbb{R}^N$. Then we can identify the finite dimensional parameters γ_k and $\gamma\in\Gamma^N$ with the vectors $\mathbf{c}^k$ and $\mathbf{c}\in\mathbb{R}^N$, so we simply write $x(t_i,\mathbf{c}^k)$ and $J^{k,N}(\mathbf{c})$ instead of $x(t_i,\gamma_k)$ and $J^{k,N}(\gamma)$. Then we have

$$\begin{aligned} J^{k,N}(\mathbf{c}) &= \sum_{i=0}^{l}\Big(x(t_i,\mathbf{c}^k) + D_2x(t_i,\mathbf{c}^k)\sum_{j=1}^{N}(c_j-c_j^k)\chi_j^N - X_i\Big)^2 \\ &= \sum_{i=0}^{l}\Big(x(t_i,\mathbf{c}^k) - X_i + \sum_{j=1}^{N}(c_j-c_j^k)D_2x(t_i,\mathbf{c}^k)\chi_j^N\Big)^2. \end{aligned}$$

To find the minimizer of $J^{k,N}(\mathbf{c})$ first consider

$$\frac{\partial}{\partial c_p}J^{k,N}(\mathbf{c}) = 2\sum_{i=0}^{l}\Big(x(t_i,\mathbf{c}^k) - X_i + \sum_{j=1}^{N}(c_j-c_j^k)D_2x(t_i,\mathbf{c}^k)\chi_j^N\Big)D_2x(t_i,\mathbf{c}^k)\chi_p^N.$$

We introduce the N-dimensional vectors

$$\mathbf{m}(t_i,\mathbf{c}^k) := \Big(D_2x(t_i,\mathbf{c}^k)\chi_1^N,\ldots,D_2x(t_i,\mathbf{c}^k)\chi_N^N\Big)^T, \tag{3.1}$$

$$\mathbf{b}(\mathbf{c}^k) := \sum_{i=0}^{l}\mathbf{m}(t_i,\mathbf{c}^k)(x(t_i,\mathbf{c}^k)-X_i) \tag{3.2}$$

and the $N\times N$ matrix

$$\mathbf{D}(\mathbf{c}^k) := \sum_{i=0}^{l}\mathbf{m}(t_i,\mathbf{c}^k)\mathbf{m}^T(t_i,\mathbf{c}^k). \tag{3.3}$$

Then $\frac{\partial}{\partial c_p}J^{k,N}(\mathbf{c}) = 0$ for $p=1,\ldots,N$, if and only if

$$\mathbf{D}(\mathbf{c}^k)(\mathbf{c}-\mathbf{c}^k) = -\mathbf{b}(\mathbf{c}^k). \tag{3.4}$$

We note that the Hessian of $J^{k,N}(\mathbf{c})$ is $2\mathbf{D}(\mathbf{c}^k)$.

Lemma 3.1. $\mathbf{D}(\mathbf{c}^k)$ *is a positive semi-definite $N\times N$ matrix, and it is positive definite if and only if there is no $\mathbf{u}\in\mathbb{R}^N$ such that $\mathbf{u}\neq\mathbf{0}$ and $\mathbf{u}\perp\mathbf{m}(t_i,\mathbf{c}^k)$ for $i=0,\ldots,\ell$.*

Proof. Let $\mathbf{u} \in \mathbb{R}^N$ and consider

$$\mathbf{u}^T \mathbf{D}(\mathbf{c}^k)\mathbf{u} = \sum_{i=0}^{l} \mathbf{u}^T \mathbf{m}(t_i, \mathbf{c}^k)\mathbf{m}^T(t_i, \mathbf{c}^k)\mathbf{u} = \sum_{i=0}^{l} \Big(\mathbf{m}^T(t_i, \mathbf{c}^k)\mathbf{u}\Big)^T \mathbf{m}^T(t_i, \mathbf{c}^k)\mathbf{u} \geq 0,$$

which yields the statement of the lemma. □

Suppose $\mathbf{c}^0 \in \mathbb{R}^N$ is given, and $\mathbf{D}(\mathbf{c}^k)$ is invertible for all $k = 0, 1, \ldots$. Then we define the QL method by the iteration

$$\mathbf{c}^{k+1} = \mathbf{c}^k - \mathbf{D}^{-1}(\mathbf{c}^k)\mathbf{b}(\mathbf{c}^k), \qquad k = 0, 1, \ldots. \tag{3.5}$$

Lemma 3.1 and the previous calculation imply that $\mathbf{c}^{k+1}$ is the unique minimizer of $J^{k,N}(\mathbf{c})$.

This is the same scheme that was used in [5] and [6] except that there the parameter space was finite dimensional, and the set $\{\chi_1^N, \ldots, \chi_N^N\}$ was the canonical basis of $\mathbb{R}^N$. In our examples the parameter space will be the space of Lipschitz continuous functions, and therefore $D_2x(t_i, \mathbf{c}^k)$ is a linear functional defined on the space of $W^{1,\infty}$-functions, and $D_2x(t_i, \mathbf{c}^k)\chi_j^N$ denotes the value of the linear functional applied to the function χ_j^N. For an alternative derivation of the QL method for ODEs with finite dimensional parameters we refer to [3].

4. **Convergence results.**

In this section we show the local convergence of the scheme (3.5) supposing the existence of an exact fit solution of the parameter estimation problem $\mathcal{P}$. We assume

(B1) $\Gamma^N \subset \Gamma$ is a finite dimensional subspace for all $N \in \mathbb{N}$;
(B2) there exists $\gamma^* \in \Gamma$, for which $J(\gamma^*) = 0$.

The next theorem studies the convergence of the QL scheme (3.5) in the case when $\gamma^* \in \Gamma^N$ for some $N \in \mathbb{N}$.

Definition 4.1. *We say that the sequence $\mathbf{c}^k \in \mathbb{R}^N$ converges to $\mathbf{c}^* \in \mathbb{R}^N$ superlinearly if there exists a sequence $\varepsilon_k \geq 0$ such that $\varepsilon_k \to 0$ as $k \to \infty$, and*

$$|\mathbf{c}^{k+1} - \mathbf{c}^*| \leq \varepsilon_k |\mathbf{c}^k - \mathbf{c}^*|, \qquad k \in \mathbb{N}.$$

An iterative method is locally convergent to $\mathbf{c}^*$ if there exists a neigborhood V of $\mathbf{c}^*$ with the property that for all initial value from V the method converges to $\mathbf{c}^*$.

Theorem 4.2. *Assume (A1) (i)–(iii), (A2) (i)–(iii) and (B1)–(B2). Suppose $\gamma^* \in P_1$, and suppose $\gamma^* = \sum_{j=1}^N c_j^* \chi_j^N \in \Gamma^N$ for some $N \in \mathbb{N}$, and $\mathbf{D}(\mathbf{c}^*)$ is invertible where $\mathbf{c}^* := (c_1^*, \ldots, c_N^*)^T$. Then for this N the QL sequence (3.5) is locally superlinearly convergent to $\mathbf{c}^*$.*

Proof. Since P_1 is a open set, it follows from Theorem 2.2 that there exists $\delta_1 > 0$ such that $D_2x(t, \gamma) \in \mathcal{L}(\Gamma, \mathbb{R})$ exists and it is continuous for $t \in [0, \alpha]$ and $\gamma \in \mathcal{B}_\Gamma(\gamma^*; \delta_1)$. Then there exists $\delta_2 > 0$ such that for $|\mathbf{c} - \mathbf{c}^*| < \delta_2$ the corresponding parameter $\gamma = \sum_{j=1}^N c_j \chi_j^N \in \mathcal{B}_\Gamma(\gamma^*; \delta_1)$. Hence $\mathbf{D}(\mathbf{c})$ is well-defined and continuous on $\mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*; \delta_2)$. Since $\mathbf{D}(\mathbf{c})$ is invertible at $\mathbf{c}^*$ and continuous, there exist $0 < \delta_3 \leq \delta_2$ and $d > 0$ such that $\mathbf{D}(\mathbf{c})$ is invertible and satisfies

$$\Big|\mathbf{D}^{-1}(\mathbf{c})\Big| \leq d, \qquad \text{for } \mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*; \delta_3).$$

Then the function

$$\mathbf{g}\colon \mathbb{R}^N \supset \mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\, \delta_3) \to \mathbb{R}^N, \qquad \mathbf{g}(\mathbf{c}) := \mathbf{c} - \mathbf{D}^{-1}(\mathbf{c})\mathbf{b}(\mathbf{c})$$

is well-defined. Consider

$$\begin{aligned} g(\mathbf{c}) - \mathbf{c}^* &= \mathbf{c} - \mathbf{c}^* - \mathbf{D}^{-1}(\mathbf{c})\mathbf{b}(\mathbf{c}) \\ &= \mathbf{D}^{-1}(\mathbf{c})\Big(\mathbf{D}(\mathbf{c})(\mathbf{c}-\mathbf{c}^*) - \mathbf{b}(\mathbf{c})\Big) \\ &= \mathbf{D}^{-1}(\mathbf{c})\sum_{i=0}^{l} \mathbf{m}(t_i,\mathbf{c})\Big(\mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\mathbf{c}^*) - (x(t_i,\mathbf{c}) - X_i)\Big). \end{aligned} \tag{4.1}$$

Now using the exact fit-to-data assumption, $\mathbf{c}^*$ satisfies $x(t_i,\mathbf{c}^*) = X_i$ for $i = 1,\ldots,N$, hence (4.1) yields

$$g(\mathbf{c}) - \mathbf{c}^* = -\mathbf{D}^{-1}(\mathbf{c})\sum_{i=0}^{l} \mathbf{m}(t_i,\mathbf{c})\Big(x(t_i,\mathbf{c}) - x(t_i,\mathbf{c}^*) - \mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\mathbf{c}^*)\Big). \tag{4.2}$$

It follows from (2.8) that

$$|D_2x(t_i,\mathbf{c})\chi_j^N| \le N_1|\chi_j^N|_\Gamma \qquad \text{for } i = 0,\ldots,\ell,\ \mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\,\delta_3),\ \text{and } j = 1,\ldots,N.$$

Then there exists $m_0 > 0$ such that

$$|\mathbf{m}(t_i,\mathbf{c})| \le m_0, \qquad i = 0,\ldots,\ell, \quad \mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\,\delta_3). \tag{4.3}$$

Hence (4.2) implies

$$|g(\mathbf{c}) - \mathbf{c}^*| \le dm_0 \sum_{i=0}^{l}\Big|x(t_i,\mathbf{c}) - x(t_i,\mathbf{c}^*) - \mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\mathbf{c}^*)\Big|, \qquad \mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\,\delta_3).$$

We have

$$\mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\mathbf{c}^*) = D_2x(t_i,\gamma)(\gamma-\gamma^*),$$

where $\gamma := \sum_{j=1}^N c_j\chi_j^N$ and $\gamma^* := \sum_{j=1}^N c_j^*\chi_j^N$. Therefore

$$\begin{aligned} &x(t_i,\mathbf{c}) - x(t_i,\mathbf{c}^*) - \mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\mathbf{c}^*) \\ &\quad= D_2x(t_i,\gamma^*)(\gamma-\gamma^*) - D_2x(t_i,\gamma)(\gamma-\gamma^*) + \omega(t_i,\gamma^*,\gamma), \end{aligned} \tag{4.4}$$

where

$$\omega(t_i,\gamma^*,\gamma) := x(t_i,\gamma) - x(t_i,\gamma^*) - D_2x(t_i,\gamma^*)(\gamma-\gamma^*) \tag{4.5}$$

satisfies

$$\lim_{\gamma\to\gamma^*}\frac{|\omega(t_i,\gamma^*,\gamma)|}{|\gamma-\gamma^*|_\Gamma} = 0, \qquad i = 0,\ldots,\ell.$$

Define the vector norm on $\mathbb{R}^N$ by

$$\|\mathbf{c}\| := \Big|\sum_{j=1}^N c_j\chi_j^N\Big|_\Gamma = |\gamma|_\Gamma, \qquad \mathbf{c} \in \mathbb{R}^N.$$

Since all vector norms on $\mathbb{R}^N$ are equivalent, there exist positive constants C_1 and C_1^* such that $C_1^*|\mathbf{c}| \le \|\mathbf{c}\| = |\gamma|_\Gamma \le C_1|\mathbf{c}|$ for all $\mathbf{c} \in \mathbb{R}^N$. Then we have

$$\lim_{\mathbf{c}\to\mathbf{c}^*}\frac{|\omega(t_i,\gamma^*,\gamma)|}{|\mathbf{c}-\mathbf{c}^*|} = \lim_{\mathbf{c}\to\mathbf{c}^*}\frac{|\omega(t_i,\gamma^*,\gamma)|}{|\gamma-\gamma^*|_\Gamma}\frac{\|\mathbf{c}-\mathbf{c}^*\|}{|\mathbf{c}-\mathbf{c}^*|} \le C_1\lim_{\gamma\to\gamma^*}\frac{|\omega(t_i,\gamma^*,\gamma)|}{|\gamma-\gamma^*|_\Gamma} = 0.$$

Hence (4.4) yields

$$\begin{aligned} |g(\mathbf{c})-\mathbf{c}^*| &\le dm_0 \sum_{i=0}^{l} \Big| x(t_i,\mathbf{c}) - x(t_i,\mathbf{c}^*) - \mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\mathbf{c}^*)\Big| \\ &\le w(\mathbf{c}^*,\mathbf{c})|\mathbf{c}-\mathbf{c}^*|, \qquad \mathbf{c}\in\mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\,\delta_3), \end{aligned} \tag{4.6}$$

where

$$w(\mathbf{c}^*,\mathbf{c}) := dm_0 \sum_{i=0}^{l} \Big(C_1 |D_2 x(t_i,\gamma^*) - D_2 x(t_i,\gamma)|_{\mathcal{L}(\Gamma,\,\mathbb{R})} + \frac{|\omega(t_i,\gamma^*,\gamma)|}{|\mathbf{c}-\mathbf{c}^*|}\Big) \tag{4.7}$$

satisfies

$$\lim_{\mathbf{c}\to\mathbf{c}^*} w(\mathbf{c}^*,\mathbf{c}) = 0. \tag{4.8}$$

Hence for every $0<\nu<1$ there exists $0<\delta_4\le\delta_3$ such that $|w(\mathbf{c}^*,\mathbf{c})|\le\nu$ for $\mathbf{c}\in\mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\,\delta_4)$. Then the convergence of the sequence (3.5) follows from (4.6) for all $\mathbf{c}^0\in\mathcal{B}_{\mathbb{R}^N}(\mathbf{c}^*;\,\delta_4)$, and the superlinear speed of the convergence follows from (4.6) and (4.8). □

Next we study the case when γ^* does not belong to Γ^N for any N, but we assume that for each N the cost function J restricted to the finite dimensional parameter set Γ^N has a local infimum at $\overline{\gamma}_N\in\Gamma^N$. Then

$$J'(\overline{\gamma}_N)\chi_j^N = 2\sum_{i=0}^{\ell}(x(t_i,\overline{\gamma}_N)-X_i)D_2x(t_i,\overline{\gamma}_N)\chi_j^N = 0, \qquad j=1,\ldots,N. \tag{4.9}$$

We assume also that

(B3) for each $N\in\mathbb{N}$ the basis functions $\chi_j^N := (\chi_j^{\varphi,N},\chi_j^{\theta,N},\chi_j^{\xi,N})$ satisfy $\chi_j^{\varphi,N}\in PW^{2,\infty}$ for $j=1,\ldots,N$, and there exist mesh points $-r<t_1<\cdots<t_m<0$, where $m=m(N)$, such that $\dot\chi_j^{\varphi,N}$ and $\ddot\chi_j^{\varphi,N}$ have points of discontinuity only at t_i for all $j=1,\ldots,N$;

(B4) for each $N\in\mathbb{N}$ the fixed basis functions in Γ^N satisfy $\sum_{j=1}^N|\chi_j^N|_\Gamma\le 1$;

(B5) for each $N\in\mathbb{N}$ the cost function J restricted to the finite dimensional parameter set Γ^N has a local infimum at $\overline{\gamma}_N\in\Gamma^N$.

For the rest of this section, for simplicity, we use the 1-norm on $\mathbb{R}^N$, i.e., $|\mathbf{c}|_1 := \sum_{j=1}^N|c_j|$. The corresponding induced matrix norm on $\mathbb{R}^{N\times N}$ is denoted also by $|\cdot|_1$.

Theorem 4.3. *Assume (A1) (i)–(v), (A2) (i)–(vi), and (B1)–(B5). Suppose γ^* in (B2) satisfies $\gamma^*\in P_1$. Let $\delta^*>0$ be defined by Lemma 2.6, for a fixed $N\in\mathbb{N}$ let $\overline{\gamma}_N := \sum_{j=1}^N \overline{c}_j\chi_j^N$ be defined by (B5), $\overline{\mathbf{c}}^N := (\overline{c}_1,\ldots,\overline{c}_N)^T$, $m=m(N)$ and $\chi_j^{\varphi,N}$ $(j=1,\ldots,N)$ be defined by (B3), let*

$$K := \max\Big\{|\overline{\mathbf{c}}^N|_1+\delta^*,\ (|\overline{\mathbf{c}}^N|_1+\delta^*)\max_{j=1,\ldots,N}|\ddot\chi_j^{\varphi,N}|_{L^\infty}\Big\},$$

and let $N_3=N_3(\gamma^*,\delta^*,m,K)$ be defined by Lemma 2.6. Then if $\overline{\gamma}_N\in\mathcal{B}_\Gamma(\gamma^*;\,\delta^*)$, the matrix $\mathbf{D}(\overline{\mathbf{c}}^N)$ exists, it is invertible and satisfies

$$|\mathbf{D}^{-1}(\overline{\mathbf{c}}^N)|_1 N_3 \sum_{i=0}^{\ell}|x(t_i,\overline{\mathbf{c}}^N)-X_i| < 1.$$

Then for this fixed N the QL sequence (3.5) is locally convergent to $\overline{\mathbf{c}}^N$.

Proof. Througout this proof we associate to the vectors $\mathbf{c} := (c_1, \dots, c_N)^T \in \mathbb{R}^N$ and $\overline{\mathbf{c}}^N := (\overline{c}_1, \dots, \overline{c}_N)^T \in \mathbb{R}^N$ the parameters $\gamma_{\mathbf{c}} := \sum_{j=1}^N c_j \chi_j^N \in \Gamma^N$ and $\overline{\gamma}_N := \sum_{j=1}^N \overline{c}_j \chi_j^N \in \Gamma^N$, respectively.

We have by (B4) that $|\chi_j^N|_\Gamma \le 1$ for all $j = 1, \dots, N$, hence

$$|\gamma_{\mathbf{c}}|_\Gamma \le \sum_{j=1}^N |c_i||\chi_j^N|_\Gamma \le |\mathbf{c}|_1, \qquad \mathbf{c} \in \mathbb{R}^N. \tag{4.10}$$

As in the proof of Theorem 4.2, let δ_1 be such that $D_2 x(t,\gamma) \in \mathcal{L}(\Gamma, \mathbb{R})$ exists and it is continuous for $t \in [0,\alpha]$ and $\gamma \in \mathcal{B}_\Gamma(\gamma^*; \delta_1)$. Let $\delta^* > 0$ be defined by Lemma 2.6, and suppose that $\overline{\gamma}_N := \sum_{j=1}^N \overline{c}_j \chi_j^N \in \mathcal{B}_\Gamma(\gamma^*; \delta^*)$. Let $\delta_2 > 0$ be such that $\mathcal{B}_\Gamma(\overline{\gamma}_N; \delta_2) \subset \mathcal{B}_\Gamma(\gamma^*; \delta^*)$. Then (4.10) implies that $\gamma_{\mathbf{c}} \in \mathcal{B}_\Gamma(\overline{\gamma}_N; \delta_2)$ for $\mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta_2\big)$.

We use the notation $\gamma_{\mathbf{c}} = (\varphi_{\mathbf{c}}, \theta_{\mathbf{c}}, \xi_{\mathbf{c}}) \in \Gamma^N$. Then

$$|\varphi_{\mathbf{c}}|_{W^{1,\infty}} \le |\gamma_{\mathbf{c}}|_\Gamma \le |\mathbf{c}|_1 \le |\overline{\mathbf{c}}^N|_1 + \delta_2, \qquad \mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta_2\big).$$

It follows from assumption (B3) that $\chi_j^{\varphi,N} \in PW^{2,\infty}$, so

$$|\ddot{\varphi}_{\mathbf{c}}|_{L^\infty} \le \sum_{j=1}^N |c_i||\ddot{\chi}_j^{\varphi,N}|_{L^\infty} \le |\mathbf{c}|_1 \max_{j=1,\dots,N} |\ddot{\chi}_j^{\varphi,N}|_{L^\infty},$$

and therefore $|\varphi_{\mathbf{c}}|_{PW^{2,\infty}} \le K$ for $\mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta_2\big)$.

Let $\delta > 0$ corresponding to $\overline{\gamma}_N \in \mathcal{B}_\Gamma(\gamma^*; \delta^*)$, m and K be defined by Lemma 2.6. Then $\mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta\big)$ implies $\gamma_{\mathbf{c}} \in \mathcal{B}_\Gamma(\overline{\gamma}_N; \delta)$ using (4.10). For every d satisfying

$$|\mathbf{D}(\overline{\mathbf{c}}^N)|_1 N_3 \sum_{i=0}^{\ell} |x(t_i, \overline{\mathbf{c}}^N) - X_i| < d N_3 \sum_{i=0}^{\ell} |x(t_i, \overline{\mathbf{c}}^N) - X_i| < 1 \tag{4.11}$$

there exists $0 < \delta_3 \le \delta$ such that $\mathbf{D}(\mathbf{c})$ exists and is invertible for $\mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta_3\big)$, and $|\mathbf{D}^{-1}(\mathbf{c})| \le d$ for $\mathbf{c} \in \mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta_3\big)$.

Then the function $\mathbf{g}(c) := \mathbf{c} - \mathbf{D}^{-1}(\mathbf{c})\mathbf{b}(\mathbf{c})$ is well-defined on $\mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N; \delta_3\big)$, and similarly to (4.1) it satisfies

$$g(\mathbf{c}) - \overline{\mathbf{c}}^N = \Big(\mathbf{D}(\mathbf{c})\Big)^{-1} \sum_{i=0}^{l} \mathbf{m}(t_i, \mathbf{c}) \Big(\mathbf{m}^T(t_i, \mathbf{c})(\mathbf{c} - \overline{\mathbf{c}}^N) - (x(t_i, \mathbf{c}) - X_i)\Big). \tag{4.12}$$

It follows from (4.9) that

$$\sum_{i=0}^{\ell} (x(t_i, \overline{\mathbf{c}}^N) - X_i) \mathbf{m}(t_i, \overline{\mathbf{c}}^N) = \mathbf{0},$$

hence combining the above with (4.12) gives

$$\begin{aligned} g(\mathbf{c}) - \overline{\mathbf{c}}^N \quad &= \quad \Big(\mathbf{D}(\mathbf{c})\Big)^{-1} \sum_{i=0}^{l} \mathbf{m}(t_i, \mathbf{c}) \Big(\mathbf{m}^T(t_i, \mathbf{c})(\mathbf{c} - \overline{\mathbf{c}}^N) - (x(t_i, \mathbf{c}) - x(t_i, \overline{\mathbf{c}}^N)\Big) \\ &\quad - \Big(\mathbf{D}(\mathbf{c})\Big)^{-1} \sum_{i=0}^{l} \Big(\mathbf{m}(t_i, \mathbf{c}) - \mathbf{m}(t_i, \overline{\mathbf{c}}^N)\Big)(x(t_i, \overline{\mathbf{c}}^N) - X_i). \end{aligned} \tag{4.13}$$

Then using (2.15) and (B4) we get

$$\begin{aligned}
|\mathbf{m}(t_i,\mathbf{c})-\mathbf{m}(t_i,\overline{\mathbf{c}}^N)|_1 &= \sum_{j=1}^{N}|D_2x(t_i,\gamma_{\mathbf{c}})\chi_j^N - D_2x(t_i,\overline{\gamma}^N)\chi_j^N| \\
&\le N_3|\gamma_{\mathbf{c}}-\overline{\gamma}^N|_\Gamma \sum_{j=1}^{N}|\chi_j^N|_\Gamma \\
&\le N_3|\mathbf{c}-\overline{\mathbf{c}}^N|_1, \quad i=0,\ldots,\ell,\ \mathbf{c}\in\mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N;\,\delta_3\big). \qquad (4.14)
\end{aligned}$$

Let m_0, ω and w be defined by (4.3), (4.5) and (4.7), respectively. Then (4.6), (4.13) and (4.14) yield

$$\begin{aligned}
|g(\mathbf{c})-\overline{\mathbf{c}}^N|_1 &\le dm_0\sum_{i=0}^{l}\Big|\mathbf{m}^T(t_i,\mathbf{c})(\mathbf{c}-\overline{\mathbf{c}}^N)-(x(t_i,\mathbf{c})-x(t_i,\overline{\mathbf{c}}^N))\Big|_1 \\
&\quad +d\sum_{i=0}^{l}|\mathbf{m}(t_i,\mathbf{c})-\mathbf{m}(t_i,\overline{\mathbf{c}}^N)|_1|x(t_i,\overline{\mathbf{c}}^N)-X_i| \\
&\le (w(\overline{\mathbf{c}}^N,\mathbf{c})+A_N)|\mathbf{c}-\overline{\mathbf{c}}^N|_1, \qquad \mathbf{c}\in\mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N;\,\delta_3\big), \qquad (4.15)
\end{aligned}$$

where by (4.11)

$$A_N := dN_3\sum_{i=0}^{l}|x(t_i,\overline{\mathbf{c}}^N)-X_i|<1.$$

Let ν be such that $A_N<\nu<1$. Then (4.8) yields that there exists $0<\delta_4\le\delta_3$ such that $0\le w(\overline{\mathbf{c}}^N,\mathbf{c})<\nu-A_N$ for $\mathbf{c}\in\mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N;\,\delta_4\big)$. Therefore (4.15) gives

$$|\mathbf{c}^{k+1}-\overline{\mathbf{c}}^N|_1\le\nu|\mathbf{c}^k-\overline{\mathbf{c}}^N|_1, \qquad \mathbf{c}^0\in\mathcal{B}_{\mathbb{R}^N}\big(\overline{\mathbf{c}}^N;\,\delta_4\big),$$

which proves the local convergence of (3.5) to $\overline{\mathbf{c}}^N$. □

5. **Numerical examples.** In all the numerical examples we present below only one component of the parameter vector (φ,θ,ξ) is considered to be unknown, the other two components will be given. So the parameter set Γ will be identified with either $W^{1,\infty}$, Θ or Ξ. Also, θ and ξ below will be coefficient functions in the equations, so we will use $W^{1,\infty}([0,\alpha],\mathbb{R})$ as the parameter set for Θ or Ξ. In all these three cases we approximate the functions of $W^{1,\infty}$ or $W^{1,\infty}([0,\alpha],\mathbb{R})$ by linear splines. Hence in the examples we define Γ^N as the space of linear spline functions with equally distant node points $\nu_1,\nu_2,\ldots,\nu_N$ of the domain $[-r,0]$ or $[0,\alpha]$. Let $\{\lambda_1^N,\ldots,\lambda_N^N\}$ be the usual "hat" functions corresponding to the mesh $\{\nu_1,\ldots,\nu_N\}$ satisfying $\lambda_i^N(\nu_j)=0$ if $i\ne j$, and $\lambda_i^N(\nu_i)=1$. Then the basis of Γ^N will be the scaled "hat" functions $\{\chi_1^N,\ldots,\chi_N^N\}$ defined by $\chi_i^N(t):=\frac{1}{N|\lambda_i^N|_{W^{1,\infty}}}\lambda_i^N$ for $i=1,\ldots,N$. Then Γ^N and $\{\chi_1^N,\ldots,\chi_N^N\}$ satisfy assumptions (B1), (B3) and (B4).

Example 5.1. Consider the scalar SD-DDE

$$\dot{x}(t) = \theta(t)x\Big(t-\xi^2(t)x^2(t)-1\Big), \qquad t\in[0,3], \qquad (5.1)$$

$$x(t) = \varphi(t), \qquad t\in[-r,0]. \qquad (5.2)$$

If we take

$$\xi(t) := \frac{20}{(t+4)^2}, \qquad \theta(t) := \frac{2t+8}{(t+2)^2} \quad \text{and} \quad \varphi(t) := \frac{1}{20}(t+4)^2 \tag{5.3}$$

as the parameters in (5.1)-(5.2), then the solution of the corresponding IVP (5.1)–(5.2) is

$$x(t) = \frac{1}{20}(t+4)^2. \tag{5.4}$$

Note that along with the "true" solution (5.4), the time lag function is $t - x^2(t)\xi^2(t) - 1 = t - 2$, so $r \geq 2$ is needed in (5.2) to generate solution (5.4).

We used the function (5.4) to generate measurements at the points $t_i = 0.2i$, $i = 0, 1, \ldots, 15$. In this example let ξ and φ be defined by (5.3), and consider θ as an unknown parameter in the equation. The derivative of the solution $x(t, \theta)$ of the IVP (5.1)–(5.2) with respect to θ applied to a fixed function $h \in W^{1,\infty}([0,3], \mathbb{R})$ is denoted by $z(t) := z(t, \theta, h) = D_2 x(t, \theta)h$, and it satisfies the variational equation

$$\begin{aligned}
\dot{z}(t) &= \theta(t)\Big[-\dot{x}\Big(t - \xi^2(t)x^2(t) - 1\Big)\xi^2(t)2x(t)z(t) + z\Big(t - \xi^2(t)x^2(t) - 1\Big)\Big] \\
&\quad + h(t)x\Big(t - \xi^2(t)x^2(t) - 1\Big), \qquad t \in [0,3], \qquad (5.5) \\
z(t) &= 0, \qquad t \in [-2, 0], \qquad (5.6)
\end{aligned}$$

where $x(t) = x(t, \theta)$. To apply iteration (3.5) we fix N, pick an initial guess for the unknown parameter, i.e., for $\mathbf{c}^0$, and starting with $k = 0$ we have to compute $x(t_i, \mathbf{c}^k)$ and $D_2 x(t_i, \mathbf{c}^k)\chi_j^N$ for $i = 0, \ldots, \ell$ and $j = 1, \ldots, N$, since they are needed to evaluate $\mathbf{D}(\mathbf{c}^k)$ and $\mathbf{b}(\mathbf{c}^k)$. In this (and also in the next examples) we approximate $x(t_i, \mathbf{c}^k)$ and $D_2 x(t_i, \mathbf{c}^k)\chi_j^N$ by solving the IVP (5.1)-(5.2) and IVP (5.5)-(5.6) numerically with step size 0.05 by the approximation technique introduced in [12]. We note that despite of this approximation technique is only of first order, our numerical runnings show that the QL iteration using these approximate function values gives a good estimate of the true parameter value in a few steps.

First we computed iteration (3.5) starting from the constant 0 initial parameter value. The numerical results can be seen in Figures 1 and 2 using $N = 3$ and $N = 8$ dimensional linear spline approximations of the coefficient function θ. In the figures the solid curve represents the "true" parameter function θ, and the dotted curves are the spline approximations obtained by the QL sequence (3.5). We observe good approximation of the "true" parameter θ in two steps. In Tables 1 and 2 the value of the least square cost function $J(\theta^{(k)})$ at the kth iteration, and the the error of the spline iteration function at the node points $\Delta_i^{(k)} = |\theta^{(k)}(\nu_i) - \theta(\nu_i)|$ are presented.

Let $P^N f$ denote the projection of the function f to the space of N-dimensional linear spline functions (with equi-distant node points). In Figures 3 and 4 and Tables 3 and 4 the numerical results of the iteration (3.5) can be seen starting from the initial parameter guess $\theta^{(0)}(t) = P^3(4\sin 5t)$ and $\theta^{(0)}(t) = P^8(4\sin 5t)$, respectively. As in the previous running, a quick convergence is observed.

Table 1: $\theta^{(0)}(t) = 0$, $N = 3$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$
0:	13.257248	2.00000	0.89796	0.56000
1:	0.583975	0.10736	0.31157	0.41742
2:	0.000202	0.25890	0.04866	0.02411

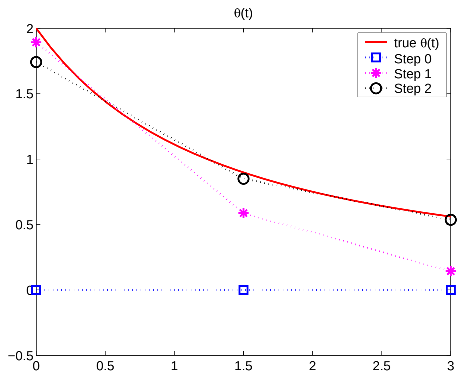

Figure 1: $\theta^{(0)}(t) = 0,\ N = 3$

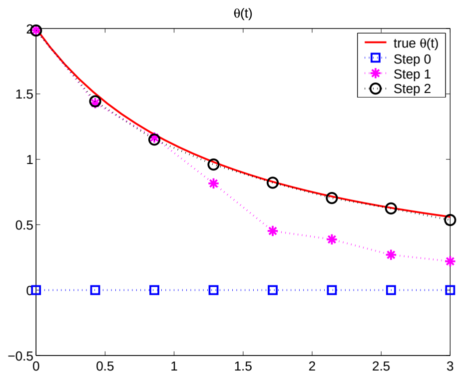

Figure 2: $\theta^{(0)}(t) = 0,\ N = 8$

Table 2: $\theta^{(0)}(t) = 0,\ N = 8$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$	$\Delta_4^{(k)}$	$\Delta_5^{(k)}$	$\Delta_6^{(k)}$	$\Delta_7^{(k)}$	$\Delta_8^{(k)}$
0:	13.25725	2.00000	1.50173	1.19000	0.97921	0.82840	0.71581	0.62891	0.56000
1:	0.57743	0.01275	0.07210	0.02331	0.16346	0.37610	0.32800	0.35868	0.33955
2:	0.00001	0.01554	0.05837	0.03913	0.01889	0.00730	0.01190	0.00464	0.02400

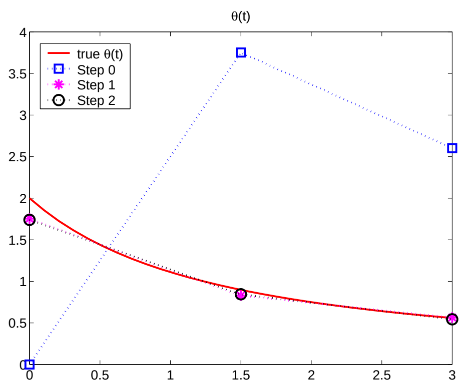

Figure 3: $\theta^{(0)}(t) = P^3(4\sin 5t),\ N = 3$

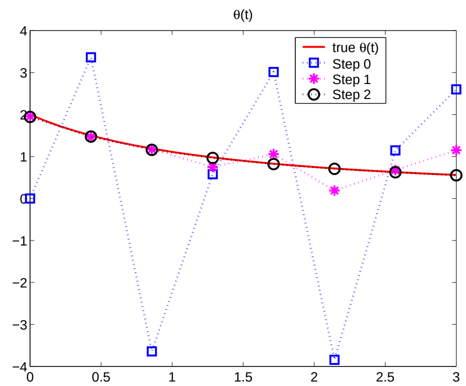

Figure 4: $\theta^{(0)}(t) = P^8(4\sin 5t),\ N = 8$

Table 3: $\theta^{(0)}(t) = P^3(4\sin 5t),\ N = 3$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$
0:	10.318073	2.00000	2.85404	2.04115
1:	0.000319	0.24502	0.06980	0.00527
2:	0.000179	0.26077	0.05294	0.01625
3:	0.000177	0.26321	0.05177	0.01668

Table 4: $\theta^{(0)}(t) = P^8(4\sin 5t),\ N = 8$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$	$\Delta_4^{(k)}$	$\Delta_5^{(k)}$	$\Delta_6^{(k)}$	$\Delta_7^{(k)}$	$\Delta_8^{(k)}$
0:	11.80723	2.00000	1.86142	4.83139	0.39971	2.18554	4.55861	0.51786	2.04115
1:	0.05504	0.04142	0.03820	0.01805	0.23000	0.22969	0.52617	0.04923	0.59118
2:	0.00000	0.05690	0.02693	0.03152	0.01420	0.00792	0.00952	0.00417	0.00684

Example 5.2. In this example we consider again the IVP (5.1)-(5.2), where now we suppose φ and θ are defined by (5.3), and we consider ξ in (5.1) as an unknown parameter function defined on the interval $[0, 3]$. We use the same measurement

generated by the "true solution" (5.4) which was used in Example 5.1. The derivative of the solution $x(t,\xi)$ of IVP (5.1)–(5.2) with respect to ξ applied to a fixed function $h \in W^{1,\infty}([0,3],\mathbb{R})$ is denoted by $z(t) := z(t,\xi,h) = D_2x(t,\xi)h$, and it satisfies the variational equation

$$\begin{aligned} \dot z(t) &= \theta(t)\Big[-\dot x\Big(t-\xi^2(t)x^2(t)-1\Big)\Big(\xi^2(t)2x(t)z(t)+2\xi(t)x^2(t)h(t)\Big) \\ &\quad +z\Big(t-\xi^2(t)x^2(t)-1\Big)\Big], \qquad t\in[0,3], \qquad (5.7)\\ z(t) &= 0, \qquad t\in[-2,0], \qquad (5.8)\end{aligned}$$

where $x(t) = x(t,\theta)$. We used the numerical solution of the IVP (5.7)-(5.8) to compute the QL sequence (3.5). We generated the sequence starting from the initial parameter value $\xi^{(0)}(t) = 1$. The first several terms of the corresponding sequence is illustrated in Figures 5 and 6 and in Tables 5 and 6 using $N = 3$ and $N = 8$ dimensional spline approximation, respectively.

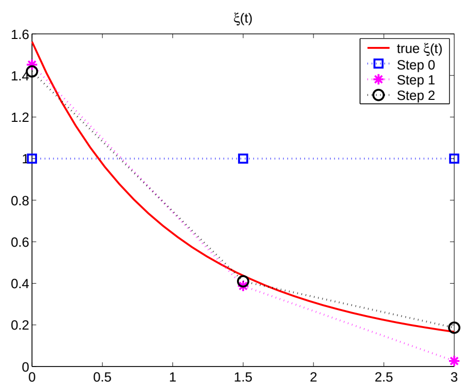

Figure 5: $\xi^{(0)}(t) = 1$, $N = 3$

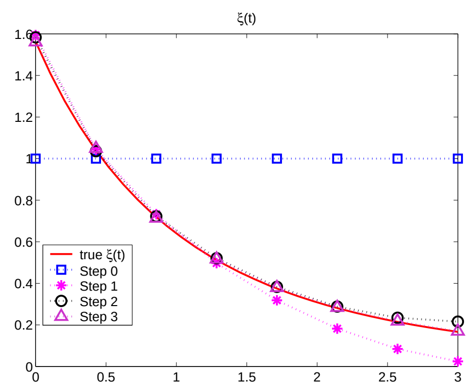

Figure 6: $\xi^{(0)}(t) = 1$, $N = 8$

Table 5: $\xi^{(0)}(t) = 1$, $N = 3$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$
0:	1.419877	0.56250	0.56287	0.83340
1:	0.080676	0.11016	0.04972	0.13968
2:	0.000964	0.14078	0.02789	0.01848
3:	0.000219	0.14846	0.02439	0.00513

Table 6: $\xi^{(0)}(t) = 1$, $N = 8$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$	$\Delta_4^{(k)}$	$\Delta_5^{(k)}$	$\Delta_6^{(k)}$	$\Delta_7^{(k)}$	$\Delta_8^{(k)}$
0:	1.419877	0.56250	0.03993	0.28132	0.48756	0.62484	0.71908	0.78550	0.83340
1:	0.078229	0.03357	0.00237	0.01607	0.01850	0.05421	0.09934	0.12863	0.14326
2:	0.001305	0.02226	0.00555	0.00493	0.00522	0.00288	0.01240	0.01409	0.06391
3:	0.000049	0.00075	0.00574	0.00230	0.00027	0.00042	0.00531	0.00153	0.00614

Example 5.3. Now consider again the IVP (5.1)-(5.2), where the coefficients θ and ξ are defined by (5.3), and in this example we consider the initial function φ as the unknown parameter in the equation. We use the same measurements that was used in Examples 5.1 and 5.2, therefore the true parameter value will be the function φ defined in (5.3).

Note that the difficulty to estimate the initial function in SD-DDEs is that the size of the initial interval depends on the solution, therefore it is not known in advance. One simple trick is to handle this difficulty numerically is to modify the initial condition in the computation of the numerical solution of (5.1). Using the measurements X_i at the time mesh points t_i and the formula of the delay function we select r so that $-r \geq \max(\xi^2(t_i)X_i^2 + 1)$, consider a function $\varphi \in W^{1,\infty}([-r,0],\mathbb{R})$, and we replace (5.2) by the initial condition

$$x(t) = \begin{cases} \varphi(t), & t \in [-r,0], \\ \varphi(-r), & t < -r. \end{cases}$$

The derivative of the solution $x(t,\varphi)$ of IVP (5.1)–(5.2) with respect to φ applied to a fixed function $h \in W^{1,\infty}([-r,0],\mathbb{R})$ is denoted by $z(t) := z(t,\varphi,h) = D_2x(t,\varphi)h$, and it satisfies the variational equation

$$\begin{aligned} \dot{z}(t) &= \theta(t)\Big[-\dot{x}\Big(t-\xi^2(t)x^2(t)-1\Big)\xi^2(t)2x(t)z(t) \\ &\quad +z\Big(t-\xi^2(t)x^2(t)-1\Big)\Big], \qquad t\in[0,3], \qquad (5.9) \\ z(t) &= h(t), \qquad t \in [-r,0], \qquad (5.10) \end{aligned}$$

where $x(t) = x(t,\theta)$. Again, in the numerical computation we replace (5.10) by

$$z(t) = \begin{cases} h(t), & t \in [-r,0], \\ h(-r), & t < -r. \end{cases}$$

In the generation of the iteration (3.5) below we used $r = 2$ and the projection of the function $\cos t$ to the space of linear spline functions as the initial parameter value. The numerical results can be seen in Figures 7 and 8 and in Tables 7 and 8 for $N = 3$ and $N = 8$, respectively. We observe quick convergence of the approximating sequences to the true parameter function φ. We note that in this example we observed convergence of the iteration scheme for picking the initial parameter value only from a small neighborhood of the true parameter. In the previous two examples the QL method was convergent in a much larger parameter region.

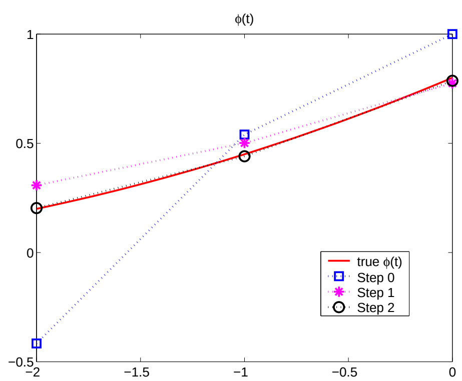

Figure 7: $\varphi^{(0)}(t) = P^3(\cos t)$, $N = 3$

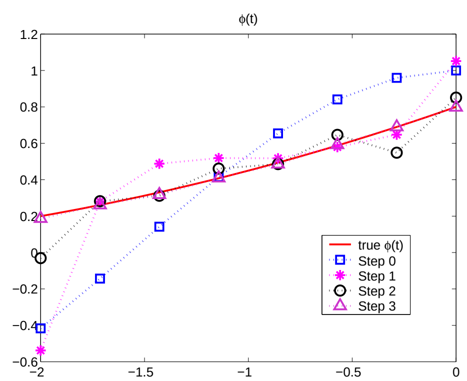

Figure 7: $\varphi^{(0)}(t) = P^8(\cos t)$, $N = 8$

Table 5: $\varphi^{(0)}(t) = P^3(\cos t)$, $N = 3$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$
0:	0.082319	0.61615	0.09030	0.20000
1:	0.108323	0.10783	0.05159	0.02523
2:	0.000084	0.00364	0.00916	0.01367
3:	0.000011	0.00592	0.01128	0.00583
4:	0.000005	0.00828	0.01205	0.00373

Table 6: $\varphi^{(0)}(t) = P^8(\cos t)$, $N = 8$

k	$J(\theta^{(k)})$	$\Delta_1^{(k)}$	$\Delta_2^{(k)}$	$\Delta_3^{(k)}$	$\Delta_4^{(k)}$	$\Delta_5^{(k)}$	$\Delta_6^{(k)}$	$\Delta_7^{(k)}$	$\Delta_8^{(k)}$
0:	0.172338	0.61615	0.40422	0.18887	0.00683	0.16072	0.25337	0.26966	0.20000
1:	0.110547	0.73788	0.01933	0.15739	0.11087	0.02379	0.00866	0.04256	0.25172
2:	0.001212	0.23078	0.02075	0.01854	0.05279	0.00820	0.05878	0.14140	0.05103
3:	0.000005	0.01346	0.00017	0.01250	0.00098	0.00847	0.00407	0.00027	0.00237

We refer to [16] for more numerical examples of the QL method (3.5) for SD-DDEs. We note that the parameter estimation problem for several classes of state-dependent and also for state-independent delay and neutral equations was studied in [1, 2, 8, 20, 21, 23, 24, 27, 32] using direct finite dimensional optimization methods. Finally note that the identifiability of parameters, i.e., the uniqueness of the parameter value which generate the same solution is an important issue in the theory of parameter estimation. It is studied for FDEs, e.g., in [30, 33], but similar studies are missing for SD-FDEs. We refer to Example 5.4 in [24], where the parameter estimation was numerically investigated in a case when the uniqueness of the parameter value failed.

REFERENCES

[1] H. T. Banks, J. A. Burns and E. M. Cliff, Parameter estimation and identification for systems with delays, SIAM J. Control and Opt., 19:6 (1981) 791–828.

[2] H. T. Banks and P. K. Daniel Lamm, Estimation of delays and other parameters in nonlinear functional differential equations, SIAM J. Control and Opt., 21:6 (1983) 895–915.

[3] H. T. Banks, G. M. Groome, Convergence theorems for parameter estimation by quasilinearization, J. Math. Anal. Appl. 42 (1973) 91–109.

[4] D. W. Brewer, The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem, SIAM. J. Math. Anal. Appl. 13:4 (1982) 607–620.

[5] D. W. Brewer, Quasi-Newton methods for parameter estimation in functional differential equations, Proc. 27th IEEE Conf. on Decision and Control, Austin, TX, (1988) 806–809.

[6] D. W. Brewer, J. A. Burns, E. M. Cliff, Parameter identification for an abstract Cauchy problem by quasilinearization, Quart. Appl. Math. 51:1 (1993) 1–22.

[7] M. Brokate, F. Colonius, Linearizing equations with state-dependent delays, Appl. Math. Optim., 21 (1990) 45–52.

[8] J. A. Burns and P. D. Hirsch, A difference equation approach to parameter estimation for differential-delay equations, Appl. Math. Comp. 7 (1980) 281–311.

[9] Y. Chen, Q. Hu, J. Wu, Second-order differentiability with respect to parameters for differential equations with adaptive delays, Front. Math. China, 5:2 (2010) 221–286.

[10] R.D. Driver, Existence theory for a delay-differential system, Contrib. Differential Equations, 1 (1961) 317–336.

[11] I. Győri and F. Hartung, On the exponential stability of a state-dependent delay equation, Acta Sci. Math. (Szeged), 66 (2000) 71–84.

[12] I. Győri, F. Hartung and J. Turi, On numerical approximations for a class of differential equations with time- and state-dependent delays, Appl. Math. Letters, 8:6 (1995) 19–24.

[13] J. K. Hale, L. A. C. Ladeira, Differentiability with respect to delays, J. Diff. Eqns., 92 (1991) 14-26.

[14] J. K. Hale and S. M. Verduyn Lunel, Introduction to Functional Differential Equations, Spingler-Verlag, New York, 1993.

[15] F. Hartung, On differentiability of solutions with respect to parameters in a class of functional differential equations, Func. Diff. Eqns., 4:1-2 (1997) 65–79.

[16] F. Hartung, Parameter estimation by quasilinearization in functional differential equations with state-dependent delays: a numerical study, Nonlinear Anal., 47:7 (2001) 4557–4566.

[17] F. Hartung, Differentiability of solutions with respect to the initial data in differential equations with state-dependent delays, J. Dynam. Differential Equations, 23:4 (2011) 843–884.

[18] F. Hartung, On differentiability of solutions with respect to parameters in neutral differential equations with state-dependent delays, Ann. Mat. Pura Appl., (2011) DOI 10.1007/s10231-011-0210-5.

[19] F. Hartung, On second-order differentiability with respect to parameters for differential equations with state-dependent delays, arXiv:1201.0269v1 [math.DS] 31 Dec 2011.

[20] F. Hartung, T. L. Herdman and J. Turi, Identifications of parameters in hereditary systems, Proceedings of ASME Fifteenth Biennial Conference on Mechanical Vibration and Noise, Boston, Massachusetts, September 1995, DE-Vol 84-3, Vol.3, Part C, 1061–1066.

[21] F. Hartung, T. L. Herdman and J. Turi, Identifications of parameters in hereditary systems: a numerical study, Proceedings of the 3rd IEEE Mediterranean Symposium on New Directions in Control and Automation, Cyprus, July 1995, 291–298.

[22] F. Hartung, T. L. Herdman, and J. Turi, On existence, uniqueness and numerical approximation for neutral equations with state-dependent delays, Appl. Numer. Math., 24 (1997) 393–409.

[23] F. Hartung, T. L. Herdman, and J. Turi, Parameter identification in classes of hereditary systems of neutral type, Appl. Math. and Comp., 89 (1998) 147–160.

[24] F. Hartung, T. L. Herdman, and J. Turi, Parameter identification in neutral functional differential equations with state-dependent delays, Nonlin. Anal., 39 (2000) 305–325.

[25] F. Hartung, T. Krisztin, H. O. Walther and J. Wu, Functional differential equations with state-dependent delays: theory and applications, in Handbook of Differential Equations: Ordinary Differential Equations, volume 3, edited by A. Canada, P. Drbek and A. Fonda, Elsevier, North-Holand, 2006, 435–545.

[26] F. Hartung and J. Turi, On differentiability of solutions with respect to parameters in state-dependent delay equations, J. Diff. Eqns. 135:2 (1997) 192–237.

[27] F. Hartung and J. Turi, Identification of Parameters in Delay Equations with State-Dependent Delays, J. Nonlinear Analysis: Theory, Methods and Applications, 29:11 (1997) 1303–1318.

[28] T. L. Herdman, P. Morin, R. D. Spies, Parameter identification for nonlinear abstract Cauchy problems using quasilinearization, J. Optim. Th. Appl. 113 (2002) 227–250.

[29] V.-M. Hokkanen and G. Morosanu, Differentiability with respect to delay, Differential and Integral Equations, 11:4 (1998) 589–603.

[30] S. M. Verduyn Lunel, Parameter identifiability of differential delay equations, Int. J. Adaptive Control Signal Processing, 15 (2001) 655–678.

[31] A. Manitius, On the optimal control of systems with a delay depending on state, control, and time. Séminaires IRIA, Analyse et Contrôle de Systèmes, IRIA, France, 1975, 149–198.

[32] K. A. Murphy, Estimation of time- and state-dependent delays and other parameters in functional differential equations, SIAM J. Appl. Math., 50:4 (1990) 972–1000.

[33] S. Nakagiri, M. Yamamoto, Identifiability of linear retarded systems in Banach spaces, Funkcialaj Ekvacioj 31 (1988) 315–329.

[34] B. Slezák, On the parameter-dependence of the solutions of functional differential equations with unbounded state-dependent delay I. The upper-semicontinuity of the resolvent function, Int. J. Qual. Theory Differential Equations Appl., 1:1 (2007) 88–114.

[35] B. Slezák, On the smooth parameter-dependence of the solutions of abstract functional differential equations with state-dependent delay, Functional Differential Equations, 17:3-4 (2010) 253–293.

[36] H. O. Walther, The solution manifold and C^1-smoothness of solution operators for differential equations with state dependent delay. J. Differential Equations 195 (2003), 46–65.

[37] H. O. Walther, Smoothness properties of semiflows for differential equations with state dependent delay. Russian, in Proceedings of the International Conference on Differential and Functional Differential Equations, Moscow, 2002, vol. 1, pp. 40–55, Moscow State Aviation Institute (MAI), Moscow 2003. English version: J. Math. Sci. 124 (2004), 5193–5207.

E-mail address: `hartung.ferenc@uni-pannon.hu`