Generalized Sensitivity Analysis for Delay Differential Equations

H. T. Banks, Danielle Robbins and Karyn L. Sutton
Center for Research in Scientific Computation
Center for Quantitative Science in Biomedicine
North Carolina State University
Raleigh, NC 27695-8212

March 8, 2012

Abstract

We present theoretical foundations for traditional sensitivity and generalized sensitivity functions for a general class of nonlinear delay differential equations. Included are theoretical results for sensitivity with respect to the delays. A brief summary of previous results along with several illustrative computational examples are also given.

1 Introduction

Delay differential equations (DDEs) have been used for a number of years to model biological, physical, and sociological processes, as well as other naturally occurring oscillatory systems. Minorsky [59] in 1942 was among the first to introduce the idea of hystero-differential equations, using these type of equations to explain self-excited oscillations arising in dynamic stabilization systems. He proposed [59, 60, 61] that some natural phenomena such as self-oscillations may be effected by the previous history of a motion or action as described by a retarded dynamical system. A retarded dynamical system is a system that describes an action that has delayed time dependence. The simplest of these physical systems are usually classified into systems with retarded damping given by

$$\ddot{x}(t) + K\dot{x}(t - \tau) + bx(t) = g(t), \quad (1.1)$$

or those with retarded restoring force described by

$$\ddot{x}(t) + K\dot{x}(t) + bx(t - \tau) = g(t), \quad (1.2)$$

where g is some external force. Specifically, Minorsky used models such as (1.1) and (1.2) to study stabilization systems in ships. It has been well understood for many years [30] that the infinite degree of the corresponding characteristic equation for a DDE such as (1.1) or (1.2) allows for an infinite number of eigenvalues for even a scalar DDE. This can promote dramatically different (from an ordinary differential equation) solution behavior such as self-excited oscillations in the solution [61]. This property of the DDE along with the widespread presence of delays in many physical and biological systems makes DDEs very important in modeling and control in these systems. Minorsky also provided insight as to the use of a nonlinear DDE to model a system with self-excited oscillations, as a linear DDE is unable to capture all of the properties of the self-excitation. Thus Minorsky lays a foundation for modeling oscillatory phenomena in general systems.

Another early contributor, Hutchinson [50], in 1948 revealed the importance of delay systems in biology and ecology. He developed a delay differential equation model, known as the delay logistic

equation, to describe the dynamics of circular causal systems. A circular causal system is any causal system (one with current solution values depending only on current or past inputs) where changes to one part of the system effects another part of the system at a different rate so that the system does not go extinct. Parasite-host interaction is an example of an ecological circular causal system; if a parasite can complete its life cycle without killing the host or drastically altering the growth of the host population, the host population will continue to exist [50, 55]. The delay in this model can represent various naturally occurring phenomena such as the gestation period in a growing population, the life cycle of a parasite, cell cycle delays, etc. Hutchinson's equation (to be used in the numerical illustrations below), its variations and other delay systems have also been used to model physiological control systems as well as numerous other biological processes [1, 23, 24, 26, 34, 35, 36, 37, 39, 40, 43, 44, 46, 47, 49, 53, 55, 56, 66]. This wide spread use of delay equations in applications has continued since the early contributions of Minorsky and Hutchinson.

In the 1970's and 80's much work was done on foundations of delay systems, in contributions both theoretical and qualitative [30, 38, 39, 48] as well as computational (see [4, 5, 6, 7, 8, 9, 20, 52, 54] and the references therein) in nature. In some of these early efforts, parameter estimation and control system questions led to the investigation of *traditional sensitivity functions (TSFs)* for delay systems. These TSFs and a more general concept of *generalized sensitivity functions (GSFs)* are the focus of our investigations in this paper. In one early paper [12], Banks, Burns and Cliff observed difficulty when estimating the delay and they suggested that this could be due to the fact that solutions of DDEs may not always be differentiable with respect to the delays; this makes estimation methods such as least squares and maximum likelihood challenging if derivative-based optimization routines are used. These authors also suggested the need for a formal theory regarding the existence of sensitivity functions with respect to the delay. Gibson and Clark [45] and Brewer [31] were among the first to treat theoretical questions of sensitivity for linear DDEs. In both contributions, these authors reformulated the delay system (as was done in the early semigroup approximation efforts of Banks, Burns and Kappel [8, 9, 20]) as an abstract system

$$\begin{aligned}\dot{z}(t) &= \mathcal{A}(q)z(t) + \mathcal{B}u(t), \quad t \geq 0 \\ z(0) &= z_0 = (\eta, \phi),\end{aligned}\tag{1.3}$$

where $(\eta, \phi) \in Z \equiv \mathbb{R}^n \times L_2(-r, 0; \mathbb{R}^n)$, $q \in Q$ and the infinitesimal generator $\mathcal{A}(q)$ is defined such that

$$\mathcal{A}(q)(\phi(0), \phi) = (L(q)\phi, \dot{\phi}).$$

Then given $t \geq 0$, $S(t; q) : Z \rightarrow Z$ is defined such that

$$S(t; q)(\eta, \phi) = (x(t; q), x_t(q))$$

where $S(t; q)$ is a strongly continuous semigroup and $x_t(\xi) = x(t + \xi)$, $-r \leq \xi \leq 0$. By defining strongly continuous solution semigroups, a well-posed problem can be formulated. While both contributions present conditions under which solutions are Frechét differentiable with respect to the parameter q , in Gibson and Clark's efforts [45] the differentiability results were obtained where the operator $\mathcal{A}(q)$ is required to be represented as a linear combination of an operator A independent of the parameter and a dependent *bounded* linear operator $A_1(q)$, i.e., $\mathcal{A}(q) = A + A_1(q)$ with A_1 bounded. Brewer [31] expands the results in [45] by considering classes of problems in which the full parameter dependent operator, $\mathcal{A}(q)$, is unbounded. In Brewer's theory this operator generates a strongly continuous semigroup, and using semigroup representation results he is able to prove the existence of Frechét derivatives with respect to the parameters for the initial value problem (1.3).

As a result of the existence of the Frechét derivatives, he is able to carefully and rigorously define sensitivity equations with respect to the parameters *including the delay* for the abstract system.

In a more recent report [2], Baker and Rihan *formally* derive sensitivity equations for delay differential equation models, as well as the equations for the sensitivity of parameter estimates with respect to observations (these latter sensitivities are what we shall discuss below as *Generalized Sensitivity Functions (GSFs)*). They consider a general nonlinear system of parameter dependent delay differential equations with parameter $p \in \mathbb{R}^L$ given by

$$\begin{aligned}\dot{x}(t, p) &= f(t, x(t), x(t - \tau), p), \quad t \geq 0, \\ x(t, p) &= \psi(t, p), \quad t \leq 0,\end{aligned}\tag{1.4}$$

and investigate methods for sensitivity functions with respect to the parameters p and delay τ .

Baker and Rihan offer an outline on how to numerically compute both TSFs and GSFs for retarded delay differential equations (as well as for neutral delay differential equations which are not discussed here in any generality). While their focus is on computational methods, they also list issues that arise when carrying out parameter estimation in DDEs. As we have already noted earlier, these include difficulty in establishing existence of the derivatives of the solution with respect to the parameters and the delays, as well as difficulty in establishing well-posedness for the derived sensitivity equations. Some of these issues are dealt with in a rigorous manner below.

Banks and Bortz [11] were among the first to consider sensitivity with respect to distributional delays. They used sensitivity analysis to show how changes in distributed parameters will effect the solutions of their nonlinear delay differential equation model for HIV progression at the cellular level where intracellular processing delays are distributed across cell populations. The models are validated with what is called *aggregate data* [7].

When deriving the sensitivity equations Banks and Bortz obtain a system of DDEs, which are assumed to be well-posed. In their discussion of well-posedness for these sensitivity equations they assume the delay distributions are differentiable and parameterizable by a mean and standard deviation. In [11] they use theoretical steps (i.e., successive approximations, fixed point theory, Lipschitz continuity, etc.) employed in [10] to prove existence and uniqueness of the resulting sensitivities and sensitivity equations. Motivated by the efforts in [11], Banks and Nguyen [25] develop a rigorous theoretical framework for sensitivity functions for general nonlinear dynamical systems in a Banach space X where the parameters μ are themselves members of another Banach space $\mathcal{M}$. In this setting they consider the sensitivity of solutions x with respect to parameters μ in the following type of abstract nonlinear ordinary differential equations

$$\begin{aligned}\dot{x}(t) &= f(t, x(t), \mu), \quad t \geq t_0 \\ x(t_0) &= x_0,\end{aligned}\tag{1.5}$$

where $f : \mathbb{R}_+ \times X \times \mathcal{M} \rightarrow X$ and $\mathcal{M}$ and X are complex Banach spaces. They establish well-posedness for (1.5), and existence of Frechét derivatives of the solution $x(t)$ with respect to the parameters μ . As a result, there is a unique solution to the corresponding sensitivity equation

$$\begin{aligned}\dot{y}(t) &= f_x(t, x(t, t_0, x_0, \mu), \mu)y(t) + f_\mu(t, x(t, t_0, x_0, \mu), \mu), \quad t \geq t_0 \\ y(t_0) &= 0,\end{aligned}\tag{1.6}$$

where $y(t) = \frac{\partial x(t)}{\partial \mu}$. In [25] Banks and Nguyen provide rigorous theoretical sensitivity results for the DDE example for HIV dynamics with measure dependent or distributional parameters given in [11]; however they only present results for the the sensitivity with respect to absolutely continuous probability distributions for the delay. In subsequent efforts [18] a rigorous theoretical foundation

is developed for sensitivity theory using *directional derivatives* where the parameter space $\mathcal{M}$ is taken as the convex metric space of probability measures (including discrete, continuous or convex combinations thereof) taken with the Prohorov metric topology [7]. Below we give new results for sensitivity with respect to discrete delays. The proofs, given in [27], while quite tedious, continue with an adaption of the well known ideas for existence and uniqueness of the Frechét derivative with respect to the delay in nonlinear DDE as employed in [11, 18, 25]. Very recent efforts, especially in areas of biology, demonstrate the continuing interest and importance of sensitivity equations in the sciences. For example Burns, Cliff, and Doughty [34] explain the use of continuous sensitivity equations for DDE models arising in a model for Chlamydia Trachomatis, while Kappel [53] discusses generalized sensitivities in dynamics of threshold-driven infections.

2 Solutions and their approximation

We first summarize some fundamental well-posedness and approximation results that have been recently developed elsewhere [7, 64]. We consider nonlinear nonautonomous dynamical systems involving delays of the general form

$$\begin{aligned}\dot{x}(t) &= G(t, x(t), x_t, x(t - \tau_1), \dots, x(t - \tau_m), \theta) + G_2(t), \quad 0 \leq t \leq T, \\ x(\xi) &= \phi(\xi), \quad -r \leq \xi \leq 0,\end{aligned}\tag{2.1}$$

where $G = G(t, \eta, \psi, y_1, \dots, y_m, \theta) : [0, T] \times X \times \mathbb{R}^{nm} \times \mathbb{R}^p \rightarrow \mathbb{R}^n$. Here $X = \mathbb{R}^n \times L_2(-r, 0; \mathbb{R}^n)$, $0 < \tau_1 < \dots < \tau_m = r$, x_t denotes the usual function $x_t(\xi) = x(t + \xi)$, $-r \leq \xi \leq 0$, and $\phi \in H^1(-r, 0)$.

The theoretical results in this manuscript will be illustrated computationally, and therefore, the solutions will need to be approximated. In order to approximate solutions, one may first convert them to an abstract evolution equation and then approximate in a space spanned by piecewise linear (or even higher order) splines (i.e., in a Galerkin approach, which is equivalent to a linear finite element approximation in partial differential equations). One is then able to numerically calculate the generalized Fourier coefficients of approximate solutions relative to the splines, and with these coefficients, recover an approximation to the solutions of (2.1).

We turn to the mathematical aspects of these nonlinear functional differential equations (FDE) systems and present an outline of the necessary mathematical and numerical foundations. First we describe the conversion of the nonlinear FDE system to an abstract evolution equation (AEE) as well as provide existence and uniqueness results for a solution to the FDE. One can use the ideas of the linear semigroup framework, in which approximation of linear delay systems has been developed, as a basis for a wide class of nonlinear delay system approximations. Details in this direction can be found in the early work [5, 6, 52] which is a direct extension of the results in [8, 9, 20] to nonlinear delay systems. We then provide a fundamental approximation framework including convergence results.

We shall make use of the following hypotheses throughout our presentation.

(H1) The function G satisfies a global Lipschitz condition:

$$\begin{aligned}|G(t, \eta, \psi, y_1, \dots, y_m, \theta) - G(t, \tilde{\eta}, \tilde{\psi}, w_1, \dots, w_m, \theta)| \\ \leq K \left(|\eta - \tilde{\eta}| + |\psi - \tilde{\psi}| + \sum_{i=1}^m |y_i - w_i| \right)\end{aligned}$$

for some fixed constant K and all $(\eta, \psi, y_1, \dots, y_m), (\tilde{\eta}, \tilde{\psi}, w_1, \dots, w_m)$ in $X \times \mathbb{R}^{nm}$ uniformly in t and θ .

(H2) The function $G : [0, T] \times X \times \mathbb{R}^{nm} \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ is differentiable.

Remark 1. If we define the function $g : [0, T] \times \mathbb{R}^n \times C(-r, 0; \mathbb{R}^n) \times \mathbb{R}^p \subset [0, T] \times X \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ given by

$$g(t, x, \theta) = g(t, \eta, \psi, \theta) = G(t, \eta, \psi, \psi(-\tau_1), \dots, \psi(-\tau_m), \theta), \quad (2.2)$$

we observe that even though G satisfies (H1), g will not satisfy a continuity hypothesis on its domain in the X norm.

Letting $z(t) = (x(t), x_t) \in X$, where the Hilbert space X has the inner product

$$\langle (\eta, \phi), (\zeta, \psi) \rangle_X = \langle \eta, \zeta \rangle_{\mathbb{R}^n} + \int_{-r}^0 \langle \phi(\xi), \psi(\xi) \rangle_{\mathbb{R}^n} d\xi, \quad (2.3)$$

we define the nonlinear operator $\mathcal{A}(t, \theta) : \mathcal{D}(\mathcal{A}) \subset X \rightarrow X$ by

$$\mathcal{D}(\mathcal{A}) \equiv \{(\psi(0), \psi) \mid \psi \in H^1(-r, 0)\}$$

$$\mathcal{A}(t, \theta)(\psi(0), \psi) = (g(t, \psi(0), \psi, \theta), D\psi)$$

where here $D\psi = \psi'$. Then the FDE (2.1) can be formulated as

$$\begin{aligned} \dot{z}(t) &= \mathcal{A}z(t) + G_2(t) \\ z(0) &= z_0, \end{aligned} \quad (2.4)$$

where $z_0 = (x_0, \phi)$ is the initial condition and $\mathcal{A} = \mathcal{A}(t, \theta)$. For notational convenience we suppress the dependence on θ in the remainder of this section.

Theorem 2. *Assume that (H1) holds and let $z(t; \phi, G_2) = (x(t; \phi, G_2), x_t(\phi, G_2))$, where x is the solution of (2.1) corresponding to $\phi \in H^1$, $G_2 \in L_2$. Then for $\zeta = (\phi(0), \phi)$, $z(t; \phi, G_2)$ is the unique solution on $[0, T]$ of*

$$z(t) = \zeta + \int_0^t [\mathcal{A}(\sigma)z(\sigma) + (G_2(\sigma), 0)] d\sigma. \quad (2.5)$$

Furthermore, $G_2 \rightarrow z(t; \phi, G_2)$ is weakly sequentially continuous from L_2 (with weak topology) to X (with strong topology).

These results can be established in one of several ways including fixed point theorem arguments or Picard iteration arguments. Either of these approaches can be used to establish existence, uniqueness and continuous dependence of the solution of (2.5). For existence, uniqueness and continuous dependence of the solution of (2.1), we note that our condition (H1) is a global version of the hypothesis of Kappel and Schappacher in [54], so that in the autonomous case their results also yield immediately the desired result for (2.1).

The uniqueness of solutions to (2.5) follows in the usual manner once we establish that $\mathcal{A}$ satisfies a dissipative inequality. Indeed, we define a weighting function w on $[-r, 0)$ by

$$w(\xi) = j \quad \text{for } \xi \in [-\tau_{m-j+1}, -\tau_{m-j}), \quad j = 1, 2, \dots, m.$$

Then we consider solutions on the space X_w which is topologically equivalent to X , with the weighted inner product

$$\langle (\eta, \phi), (\zeta, \psi) \rangle_w = \langle \eta, \zeta \rangle_{\mathbb{R}^n} + \int_{-r}^0 \langle \phi(\xi), \psi(\xi) \rangle_{\mathbb{R}^n} w(\xi) d\xi. \quad (2.6)$$

One can show without difficulty that (H1) implies the dissipative inequality (see [29, p. 71]) for the nonlinear operator $\mathcal{A}(t)$

$$\langle \mathcal{A}(t)x - \mathcal{A}(t)y, x - y \rangle_w \leq \omega \langle x - y, x - y \rangle_w \quad (2.7)$$

for all $x, y \in \mathcal{D}(\mathcal{A})$ and all t .

The system of functional differential or delay equations described can now be simulated using an algorithm first developed by Banks and Kappel for linear systems [20] and extended in [5, 6]. Solutions are approximated in a space spanned by piecewise linear splines. Thus one can numerically calculate the generalized Fourier coefficients of the approximate solution in the spline basis representation and recover an approximation to the solution of (2.1).

Define X^N to be an approximating subspace [20, 21] of X . In particular, we choose $X^N = X_1^N$ to be the piecewise linear spline subspaces of X discussed in detail in [20]. We briefly outline the results for the piecewise linear subspaces X_1^N (see Section 4 of [20]) given by

$$X_1^N = \{(\phi(0), \phi) \mid \begin{array}{l} \phi \text{ is a continuous first-order spline function} \\ \text{with knots at } t_j^N = -jr/N, j = 0, 1, \dots, N \end{array}\}.$$

A careful study of the arguments behind our presentation reveals that the approximation results given here hold for more general spline approximations. For example, if one were to treat cubic spline approximations (X_3^N of [20]), one would use the appropriate approximation analogues of Theorem 2.5 of [69] and Theorem 21 of [70] (e.g., see Theorem 4.5 of [69]). Hereafter when we write X^N the reader should understand that we mean X_1^N of [20].

Let P^N be the orthogonal projection in $\langle \cdot, \cdot \rangle_w$ of $X = X_w$ onto X^N . We define the approximating operator $\mathcal{A}^N(t) = P^N \mathcal{A}(t) P^N$ and consider the approximating equations in X^N given by

$$z^N(t) = P^N \zeta + \int_0^t [\mathcal{A}^N(\sigma) z^N(\sigma) + P^N(G_2(\sigma), 0)] d\sigma. \quad (2.8)$$

These are equivalent to

$$\dot{z}^N(t) = \mathcal{A}^N(t) z^N(t) + P^N(G_2(t), 0), \quad z^N(0) = P^N \zeta, \quad (2.9)$$

the finite dimensional system in X^N .

Define $\alpha^N(t)$ so that $x^N(t) = \hat{\beta}^N \alpha^N(t)$ for any $x^N \in X^N$. Here

$$\hat{\beta}^N = (\beta^N(0), \beta^N) \text{ where } \beta^N = (e_0^N, e_1^N, \dots, e_N^N).$$

The basis elements e_j^N 's are piecewise linear splines defined by the Kronecker symbol δ_{ij} , so

$$e_j^N(t_i) = \delta_{ij} \text{ for } i, j = 0, 1, \dots, N.$$

Then solving for $z^N(t)$ in the finite dimensional system (2.9) is equivalent to solving for $\alpha^N(t)$ in the vector system

$$\begin{aligned} \dot{\alpha}^N(t) &= A^N \alpha^N(t) + P^N G_2(t) \\ \alpha^N(0) &= \alpha_0^N, \end{aligned} \quad (2.10)$$

where $\hat{\beta}^N \alpha_0^N = P^N z_0$ and A^N is the matrix representation for $\mathcal{A}^N$. We note that having obtained $\alpha^N(t)$, the product $\hat{\beta}^N \alpha^N(t)$ converges uniformly in t to the solution $z(t) = (x(t), x_t)$ of (2.4) if we can argue the convergence $z^N(t) \rightarrow z(t)$. To do this for linear systems, one can use the Trotter-Kato theorem, involving linear semigroups. For nonlinear autonomous systems, one can invoke the use of nonlinear semigroups [52, 54].

From (2.7) and the definition of $\mathcal{A}^N$ in terms of the self-adjoint projections P^N , we have at once that under (H1) the sequence $\{\mathcal{A}^N\}$ satisfies on X a uniform dissipative inequality

$$\langle \mathcal{A}^N(t)x - \mathcal{A}^N(t)y, x - y \rangle_w \leq \omega \langle x - y, x - y \rangle_w \quad (2.11)$$

for all $x, y \in \mathcal{D}(\mathcal{A})$ and all t . Uniqueness of solutions of (2.8) then follows immediately from this inequality. Upon recognition that (2.9) is equivalent to a nonlinear ordinary differential equation in Euclidean space with the right-hand side satisfying a global Lipschitz condition, one can easily argue existence of solutions for (2.9) and hence for (2.8) on any finite interval $[0, T]$. The next theorem, which ensures that solutions of (2.9) converge to those of (2.5), along with its proof is contained in [26].

Theorem 3. *Assume (H1), (H2). Let $\zeta = (\phi(0), \phi)$, $\phi \in H^1$ and $G_2 \in L_2(0, T)$ be given, with z^N and z the corresponding solutions on $[0, T]$ of (2.9) and (2.5), respectively. Then $z^N(t) \rightarrow z(t) = (x(t; \phi, G_2), x_t(\phi, G_2))$, as $N \rightarrow \infty$, uniformly in t on $[0, T]$.*

Remark 4. One can actually obtain slightly stronger results than those given in Theorem 3. One can consider solutions of (2.5) and (2.9) corresponding to initial data $z_0 = (x_0, \phi) = \zeta$ with $x_0 \in \mathbb{R}^n$, $\phi \in L_2$ (i.e., $\zeta \in X$) and argue that the results of Theorem 3 hold also in this case.

The convergence given in Theorem 3 yields state approximation techniques for nonlinear FDE systems based on the spline methods developed in [20]. These results can be applied directly to control and identification problems, which are discussed in [5, 6].

3 Continuous dependence and differentiability

To establish continuous dependence in parameters and differentiability with respect to model parameters, initial conditions, and discrete time delays (not previously done elsewhere for general nonlinear systems to the authors' knowledge), we focus on a restricted case with nonlinear autonomous systems with one discrete delay of the form

$$\frac{dx(t)}{dt} = G(x(t), x(t - \tau), \theta), \quad t > 0 \quad (3.1)$$

$$x(\xi) = \begin{cases} \phi(\xi), & -\tau \leq \xi < 0 \\ x_0, & \xi = 0 \end{cases} \quad (3.2)$$

where $x(t) \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, and $\theta \in \mathbb{R}^p$. While we consider here the case of finite dimensional model parameters the results also hold in a more general case when parameters are distributed, and hence infinite dimensional, as presented in [17, 18]. Once established, these results allow us to study the traditional and generalized sensitivity functions, where sensitivity is considered with respect to these three quantities. We begin by considering continuous dependence of solutions $x(t)$ on model parameters θ . We note that we focus on these quantities as they are often unknown in practice and may need to be estimated from observed or experimental data. The use of sensitivity functions can aid in that endeavor.

Lemma 5. *Let $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ and for $\theta = \theta_0$, let $x(t, x_0, \phi, \tau, \theta_0)$ be a solution of (3.1) - (3.2) for $t \in [0, T]$. Assume that*

$$\lim_{\theta \rightarrow \theta_0} G(x, \tilde{x}, \theta) = G(x, \tilde{x}, \theta_0), \quad (3.3)$$

uniformly in x and $\tilde{x}$. For $(x_1, \tilde{x}_1, \theta), (x_2, \tilde{x}_2, \theta) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^p$ assume that

$$|G(x_1, \tilde{x}_1, \theta) - G(x_2, \tilde{x}_2, \theta)| \leq C_1|x_1 - x_2| + C_2|\tilde{x}_1 - \tilde{x}_2| \quad (3.4)$$

where $C_j > 0$ is a constant for $j = 1, 2$. Then the initial value problem (3.1)-(3.2) has a unique solution $x(t, x_0, \phi, \tau, \theta)$ that satisfies

$$\lim_{\theta \rightarrow \theta_0} x(t, x_0, \phi, \tau, \theta) = x(t, x_0, \phi, \tau, \theta_0), \quad t \in [0, T].$$

Next, we turn to differentiability of the general system (3.1)-(3.2) with respect to model parameters in the following theorem. The proof is excluded but can be found in [27, 64]. Also, the proof of Theorem 8 can easily be followed to give that of Theorem 6. Without further discussion, we then state Theorem 7, in which we establish differentiability of the model system with respect to the initial conditions, which is also proven in [27].

Theorem 6. Suppose that $G(x, \tilde{x}, \theta)$ has continuous Frechét derivatives $G_\theta, G_x, G_{\tilde{x}}$ such that $|G_x| \leq M_0$, $|G_{\tilde{x}}| \leq M_1$, and $|G_\theta| \leq M_2$. Then the Frechét derivative $y_1(t) = \frac{\partial x(t)}{\partial \theta} \in \mathbb{R}^{n \times p}$ exists and is the unique solution for

$$\begin{aligned} \dot{y}_1(t) &= G_x(x(t), x(t-\tau), \theta)y_1(t) + G_{\tilde{x}}(x(t), x(t-\tau), \theta)y_1(t-\tau) \\ &\quad + G_\theta(x(t), x(t-\tau), \theta), \\ y_1(t) &= 0, \quad -\tau \leq t < 0 \end{aligned}$$

Theorem 7. Suppose the function $G(x, \tilde{x}, \theta)$ of (3.1) has continuous Frechét derivatives $G_x(x, \tilde{x}, \theta)$, $G_{\tilde{x}}(x, \tilde{x}, \theta)$, with respect to x and $\tilde{x}$, with $|G_x| \leq M_0$, $|G_{\tilde{x}}| \leq M_1$. Then the Frechét derivative $y_2(t) = \frac{\partial}{\partial z_0} x(t, z_0, \theta)$ exists with $y_2(t) \in \mathcal{L}(Z, \mathbb{R}^n)$ (recall $z_0 = (x_0, \phi)$, $Z = \mathbb{R}^n \times L^2(-\tau, 0; \mathbb{R}^n)$), and satisfies the equation

$$\begin{aligned} \dot{y}_2(t)[h] &= G_x(x(t), x(t-\tau), \theta)y_2(t)[h] + G_{\tilde{x}}(x(t), x(t-\tau), \theta)y_2(t-\tau)[h], \quad t > 0 \\ y_2(\xi) &= \mathcal{I} \quad -\tau \leq \xi \leq 0, \end{aligned} \quad (3.5)$$

where $\mathcal{I} \in \mathcal{L}(Z, \mathbb{R}^n)$ is the identity.

Finally we state results for derivatives with respect to the discrete delays with proofs being given in [27].

Theorem 8. Suppose that $G(x, \tilde{x}, \theta)$ has continuous Frechét derivatives $G_x, G_{\tilde{x}}$ such that $|G_x| \leq M_0$, and $|G_{\tilde{x}}| \leq M_1$ and suppose that the solution x of (3.1)-(3.2) satisfies

$x \in H^{1,\infty}(-\tau, T; \mathbb{R}^n)$, for $0 < \tau < r$ for fixed $r > 0$. Then the Frechét derivative $y_3(t) = \frac{\partial x(t)}{\partial \tau} \in \mathbb{R}^n$ exists and is the unique solution for

$$\dot{y}_3(t) = G_x(x(t), x(t-\tau), \theta)y_3(t) + G_{\tilde{x}}(x(t), x(t-\tau), \theta)[y_3(t-\tau) - \dot{x}(t-\tau)] \quad (3.6)$$

$$y_3(\xi) = 0, \quad -\tau \leq \xi \leq 0. \quad (3.7)$$

Moreover, $\frac{\partial x(t)}{\partial \tau}$ is continuous in θ and, if $x \in C^1(-\tau, T; \mathbb{R}^n)$ it is also continuous in τ .

4 Sensitivity functions

Given the above results, especially differentiability in the quantities θ , $z_0 = (x_0, \phi)$ and τ , we are now able to use the powerful sensitivity analytic techniques in delay systems. For further simplification in the remainder of our discussions we restrict our considerations to constant function initial conditions so in $z_0 = (x_0, \phi)$ we assume $\phi(\xi) = x_0$, $-r \leq \xi \leq 0$. Traditionally, sensitivity analysis is the quantification of the effect changes in parameters have on model solutions. Traditional sensitivity functions (TSFs), which are given by,

$$\begin{aligned} y_1^k(t) &= \frac{\partial x}{\partial \theta^k} & k = 1, \dots, p \\ y_2^m(t) &= \frac{\partial x}{\partial x_0^l} & l = 1, \dots, n \\ y_3(t) &= \frac{\partial x}{\partial \tau}, \end{aligned} \tag{4.1}$$

are local in nature as they are defined by locally evaluated partial derivatives, i.e., $\frac{\partial x}{\partial \theta}(t, \bar{\theta}, \bar{x}_0, \bar{\tau})$, which gives information over specified time intervals, and at values of parameters, initial conditions and delays. Even with this limitation, these functions have been used to improve sampling in an experimental setting; specifically they can be used to guide the time at which measurements should be taken to best inform the estimation of unknown parameters [15, 16]. That is, sampling might be advisable in time intervals where, for example, $y_1^k(t)$ is large, as it indicates that the model solution $x(t)$ is sensitive to changes in the parameter θ_k . Similarly, insensitivity to a certain parameter (or unknown quantity), indicated by small or zero values of the TSF, imply that observations can not be profitably taken in that region if the goal is estimation of the parameter.

TSFs may be approximated by forward differences, but are typically found by solving the system of sensitivity equations

$$\frac{d}{dt} \frac{\partial x(t)}{\partial \theta} = \frac{\partial G}{\partial x} \frac{\partial x}{\partial \theta}(t) + \frac{\partial G}{\partial \tilde{x}} \frac{\partial x}{\partial \theta}(t - \tau) + \frac{\partial G}{\partial \theta}(t) \tag{4.2}$$

for the corresponding system

$$\begin{aligned} \frac{dx(t)}{dt} &= G(x(t), x(t - \tau), \theta), & t > 0 \\ x(\xi) &= x_0, & -r \leq \xi \leq 0, \end{aligned} \tag{4.3}$$

where the $\frac{\partial}{\partial \theta}$ and $\frac{d}{dt}$ operators have been interchanged, due to the continuity assumptions made on G and x . We note that sensitivity analysis is most efficiently carried out in two steps. Once a solution $x(t)$ corresponding to $(\bar{\theta}, \bar{x}_0, \bar{\tau})$ of the above (original delay) equation (4.3) is obtained, one uses this solution to evaluate the coefficients in system (4.2). This decoupling of the original equation and the sensitivity equation has implications when considering the sensitivity with respect to the time delay τ in one of the examples discussed below, which if solved in a coupled manner would result in a so-called neutral delay system.

Generalized sensitivity functions, first introduced by Thomseth and Cobelli [72], and further studied in a series of papers by Banks, et al., [15, 16, 19], provide a measure of how informative measurements of the output or observation variables ($f(t, q)$ defined below, which are not necessarily simply the state variables), are for the identification of unknown quantities. Notably, the functions G and h must be differentiable to construct the TSFs, and must also be sufficiently smooth to construct generalized sensitivity functions (GSFs). Before defining the GSFs we briefly outline an

inverse problem framework, not only to put our discussion in context, but also to define quantities in the definition of the GSFs.

Given a model solution $x(t)$, the sensitivity of the solution to an estimated quantity q_k (where $q = (\theta, x_0, \tau)^T$) is

$$s_k(t, q) = \frac{\partial f}{\partial q_k}(t, q) \in \mathbb{R}^m,$$

where $f(t, q) = h(t, x(t), x(t - \tau), \theta)$ are the model quantities corresponding to the observed data. Observations are typically available at discrete times, which we denote by $t_1, \dots, t_{n_d}$. The model representation of the data is then

$$f(t_j, q) = h(t_j, x(t_j), x(t_j - \tau), \theta), \quad j = 1, \dots, n_d.$$

In general, the data are not exactly $f(t_j, q)$, due to uncertainty in the measurement process, and also due to small fluctuations not explicitly included in the model. Therefore we represent the observation process Y_j at time t_j by the statistical model

$$Y_j = f(t_j, q^0) + \mathcal{E}_j, \quad j = 1, \dots, n_d, \quad (4.4)$$

where $f(t_j, q) = h(t_j, x(t_j), x(t_j - \tau), \theta)$, $q = (\theta, x_0, \tau)$, for $q \in \mathcal{Q} = \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R}^1$. Here $q^0 = (\theta^0, x_0^0, \tau^0)^T$ represents the ‘true values’ of the parameters that generates the observations $\{Y_j\}_{j=1}^{n_d}$. The existence of q^0 is commonly assumed [13], implying that (3.1) describes the biological, sociological, or physical process essentially precisely.

The observation errors $\mathcal{E}_j$ are random variables, each with unknown but assumed independent and identical probability distributions of mean zero, and constant variance σ^2 . Each data set $\{y_j\}_{j=1}^{n_d}$ is one realization of the random variable $\{Y_j\}_{j=1}^{n_d}$, and the corresponding errors are also a realization of the $\mathcal{E}_j$. Estimating unknown quantities via the minimization between the model and data assuming the statistical model (4.4) gives rise to the commonly used ordinary least squares (OLS) estimator

$$q_{OLS} = \arg \min_{q \in \mathcal{Q}} \sum_{j=1}^{n_d} |Y_j - f(t_j, q)|^2, \quad (4.5)$$

where the objective functional is minimized over an admissible parameter space $\mathcal{Q}$. Another common formulation is a weighted least squares procedure, in which the error is assumed to be proportional to the model quantity $f(t_j; q)$, i.e., relative error. For a more complete discussion of the underlying assumptions and related formulations, see [13].

The variance σ^2 of the observation error is used in the computation of standard errors, confidence intervals, etc. and also in the generalized sensitivity functions. For a given set of data, $\{y_j\}_{j=1}^{n_d}$ and parameter estimates $\hat{q}$, the (bias-adjusted) variance is estimated as

$$\hat{\sigma}^2 = \frac{1}{n - n_p} \sum_{j=1}^{n_d} |y_j - f(t_j, \hat{q})|^2 \quad (4.6)$$

for $n_p = p + n + 1$ estimated parameters, where $n_p = \dim(\mathcal{Q})$.

The generalized sensitivity functions [16, 19, 72] are defined by

$$\mathbf{gs}(t) = \int_0^t \left[F(T)^{-1} \frac{1}{\sigma^2(s)} \nabla_q f(s, q^0) \right] \cdot \nabla_q f(s, q^0) dP(s), \quad t \in [0, T], \quad (4.7)$$

for variance $\sigma^2(t)$ that may possibly be time-dependent, true parameters q^0 , some general measure P that embodies the observations, and the Fisher information matrix (FIM) F which is defined by

$$F(T) = \int_0^T \frac{1}{\sigma(t)} \nabla_q f(t, q^0) \nabla_q f(t, q^0)^T dP(t). \quad (4.8)$$

We note that the definition of the measure P affects the FIM, and it can be chosen in such a way as to optimize the information from data concerning the estimated parameters. The GSFs are cumulative functions, such that at time t_j , only the contributions of measurements up to and including those at time t_j are relevant. By the definition in (4.7), it is readily seen that the GSFs are one at the final time $\mathbf{gs}(t_{n_d}) = 1$. As discussed in [16, 72], regions over which the *sharpest change* (either increase or decrease) of the GSFs indicate regions of high information content. Decreases in the GSF corresponding to a given parameter indicate correlation between that parameter and at least one other estimated parameter. In this case, it can be seen [16] that computing the GSF for one of the correlated parameters and holding the other(s) fixed, will result in a monotonically increasing GSF. Therefore, regions over which the GSF decreases indicate that the data in that region indeed contains information concerning that parameter, but it is correlated with at least one other parameter, and simultaneous identifiability of all parameters may be difficult.

As observations are typically available at discrete time points and our discussions are in the context of parameter estimation from observed or measured data, we have included here also the definitions for the GSFs and FIM for a discrete measure $P = \sum_{j=1}^{n_d} \Delta_{t_j}$. In the discrete case, the generalized sensitivity functions are

$$\mathbf{gs}(t_j) = \sum_{i=1}^j \frac{1}{\sigma^2(t_j)} [F^{-1} \times \nabla_q f(t_i, q^0)] \cdot \nabla_q f(t_i, q^0), \quad (4.9)$$

for observation times t_j where $j = 1, \dots, n_d$. In the above definition, the discrete FIM is given by

$$F = \sum_{j=1}^{n_d} \frac{1}{\sigma^2(t_j)} \nabla_q f(t_j, q^0) \nabla_q f(t_j, q^0)^T, \quad (4.10)$$

which measures the information content of the data corresponding to the parameters. In both (4.7) and (4.9), the estimate for the variance of the observation error is used up to and including the time t_j of the observation, given by

$$\sigma^2(t_j) = \sum_{i=1}^j |y_i - f(t_i, \hat{q})|^2. \quad (4.11)$$

If the variance is assumed constant ($\sigma^2(t) \equiv \sigma^2$), one would simply calculate the estimate as in (4.6), and use that in (4.7) or (4.9).

5 Illustrative computations

To complete our presentation, we illustrate the uses of sensitivity analysis in two prominent examples of delay equations. The first example we consider is a delay version of one of the most commonly studied models, the logistic equation. This delayed logistic equation, commonly known as Hutchinson's equation mentioned above, is not only discussed in most introductory modeling courses, but is still used in research endeavors to represent growth within an environment in which

saturation is possible, but the death rate is proportional to previous population levels. The standard logistic example (without delay) has been effectively used to illustrate with simulated data the ideas of traditional and sensitivity functions and how these techniques may improve data sampling for the purpose of parameter estimation [15, 16]. Therefore, it is natural to turn to the delayed logistic equation now that we are able to study sensitivity functions in systems involving a discrete delay. Here we will also numerically generate simulated data with a known delay, and demonstrate that the estimation can be improved using insights gained from the sensitivity function solutions.

The second example we use is also an ubiquitous model, the delayed harmonic oscillator of Minorsky discussed in the Introduction. As noted there, this example arises in many physical applications where oscillatory phenomena are important.

5.1 Hutchinson equation example

In his seminal paper [50] and book [51], Hutchinson arrived at a version of the logistic equation that incorporated a delay in the carrying or death rate term,

$$\frac{dx(t)}{dt} = rx(t) \left(1 - \frac{x(t-\tau)}{K} \right). \quad (5.1)$$

The model was suggested as a possible explanation of the growth dynamics seen in *Daphnia*. This population seemed to grow exponentially at low population sizes, but it would oscillate at higher population levels. Hutchinson hypothesized that this growth was like that of the logistic model, only that the population seemed to be able to exceed its carrying capacity and perhaps it was this value that the population level was oscillating around.

The traditional sensitivity functions with respect to the model parameters r , K , initial condition x_0 , and delay τ are given by

$$\begin{aligned} \frac{\partial}{\partial r} \frac{dx(t)}{dt} &= r \left[1 - \frac{x(t-\tau)}{K} \right] \frac{\partial x(t)}{\partial r} - \frac{rx(t)}{K} \frac{\partial x(t-\tau)}{\partial r} + x(t) \left[1 - \frac{x(t-\tau)}{K} \right] \\ \frac{\partial}{\partial K} \frac{dx(t)}{dt} &= r \left[1 - \frac{x(t-\tau)}{K} \right] \frac{\partial x(t)}{\partial K} - \frac{rx(t)}{K} \frac{\partial x(t-\tau)}{\partial K} + rx(t) \left[\frac{x(t-\tau)}{K^2} \right] \\ \frac{\partial}{\partial x_0} \frac{dx(t)}{dt} &= r \left[1 - \frac{x(t-\tau)}{K} \right] \frac{\partial x(t)}{\partial x_0} - \frac{rx(t)}{K} \frac{\partial x(t-\tau)}{\partial x_0} \\ \frac{\partial}{\partial \tau} \frac{dx(t)}{dt} &= r \left[1 - \frac{x(t-\tau)}{K} \right] \frac{\partial x(t)}{\partial \tau} - \frac{rx(t)}{K} \left[\frac{\partial x(t-\tau)}{\partial \tau} - \dot{x}(t-\tau) \right]. \end{aligned}$$

By changing the order of integration, and letting $s_1(t) = \frac{\partial x(t)}{\partial r}$, $s_2(t) = \frac{\partial x(t)}{\partial K}$, $s_3(t) = \frac{\partial x(t)}{\partial x_0}$, and $s_4(t) = \frac{\partial x(t)}{\partial \tau}$, we have the system

$$\frac{\partial s_1(t)}{\partial t} = r \left[1 - \frac{x(t-\tau)}{K} \right] s_1(t) - \frac{rx(t)}{K} s_1(t-\tau) + x(t) \left[1 - \frac{x(t-\tau)}{K} \right] \quad (5.2)$$

$$\frac{\partial s_2(t)}{\partial t} = r \left[1 - \frac{x(t-\tau)}{K} \right] s_2(t) - \frac{rx(t)}{K} s_2(t-\tau) + rx(t) \left[\frac{x(t-\tau)}{K^2} \right] \quad (5.3)$$

$$\frac{\partial s_3(t)}{\partial t} = r \left[1 - \frac{x(t-\tau)}{K} \right] s_3(t) - \frac{rx(t)}{K} s_3(t-\tau) \quad (5.4)$$

$$\frac{\partial s_4(t)}{\partial t} = r \left[1 - \frac{x(t-\tau)}{K} \right] s_4(t) - \frac{rx(t)}{K} [s_4(t-\tau) - \dot{x}(t-\tau)]. \quad (5.5)$$

Figure 1: The numerical approximation to the solutions (a) for the Hutchinson equation, and the corresponding traditional (b) and generalized (c) sensitivity functions with respect to the model parameters r and K , and the constant initial value x_0 are provided here for the values $\bar{r} = 0.7$, $\bar{K} = 17.5$, and $\bar{x}_0 = 0.1$. The generalized sensitivity functions were computed with constant variance $\sigma^2 = 0.1$.

As noted earlier, we consider only the case of constant initial data, and thus we do not discuss here the Frechét derivative $y_2(t) = \frac{\partial}{\partial z_0} x(t, z_0, \theta)$ where $z_0 = (x_0, \phi)$, $Z = \mathbb{R}^n \times L^2(\tau, 0; \mathbb{R}^n)$; the results of Theorem 7 still ensure the existence and uniqueness of the solution $\frac{\partial x(t)}{\partial x_0}$ to equation (5.4), for this simpler case. The existence of unique solutions to equations (5.2) and (5.3) are guaranteed by Theorem 6, and a unique solution for equation (5.5) by Theorem 8. Note that equation (5.5) is not a neutral equation if one assumes the solution $x(t)$ (and also $x(t - \tau)$) is already computed when sensitivity analysis is done; i.e., we decouple the original equation and first solve the delay equation before computing sensitivities. Therefore, when computing sensitivities the $x(t)$ and $x(t - \tau)$ are

Figure 2: The numerical approximation for the solutions (a) to the Hutchinson equation with delay $\tau = 1$, and corresponding traditional (b) and generalized (c) sensitivity functions with respect to growth rate r , carrying capacity K , constant initial state x_0 , and delay τ , each evaluated at $(\bar{r}, \bar{K}, \bar{x}_0, \bar{\tau}) = (.7, 17.5, .1, 1)$. The generalized sensitivity functions were computed with constant variance $\sigma^2 = 0.1$.

not unknown quantities but rather an input in the traditional sensitivity functions above.

The solutions for the Hutchinson equation with no delay (i.e., the standard logistic equation), and the corresponding traditional and generalized sensitivity functions are displayed in Figure 1. In comparing panels 1(b) to 1(a), the traditional sensitivity functions with respect to the growth rate r and the initial condition x_0 suggest that the beginning growth portion of the solution is quite sensitive to both parameters. In the bottom panel 1(c), the solutions of the generalized sensitivity function suggest that the same region is informative for both parameters, but that they are correlated since one of the curves decreases as the other increases. Thus, estimating both the

Figure 3: The numerical approximation for the solutions (a) to the Hutchinson equation with delay $\tau = \frac{\pi}{2r} \approx 2.244$, and corresponding traditional (b) and generalized (c) sensitivity functions with respect to growth rate r , carrying capacity K , constant initial state x_0 , and delay τ each evaluated at $(\bar{r}, \bar{K}, \bar{x}_0, \bar{\tau}) = (.7, 17.5, .1, \frac{\pi}{2\bar{r}})$. The generalized sensitivity functions were computed with constant variance $\sigma^2 = 0.1$.

initial condition x_0 and the growth rate r simultaneously from data corresponding to this interval is likely problematic. As one would expect the solution appears to be sensitive to the carrying capacity essentially once it is approached. It is easier to see this in panel 1(c) than in 1(b), as the magnitude of the sensitivity to the other parameters (r and x_0) is significantly greater. The definition of the generalized sensitivity functions is such that their magnitude is not as varied even with respect to different quantities.

With a moderate delay, $\tau = 1$, there appears only one time interval over which the solution $x(t)$ exceeds its carrying capacity, as seen in Figure 2a. The solution then decreases to below its

Table 1: Estimation of delay τ

	$\hat{\tau}$	RSS
t_{unif}	0.9561	16.204
t_{GSF}	1.032	6.575

carrying capacity but the effect is not sufficient for the oscillations to continue, and the solution approaches its carrying capacity $x(t) \rightarrow K$ around $t = 14$. It is around the time of the solution first exceeding and then decreasing to less than the carrying capacity (approximately, the interval $t \in [8, 11]$), which can be interpreted as the effect of the delay, that the sensitivity function solutions can be interpreted to mean that the model solution $x(t)$ is sensitive to this delay τ . The solutions of the traditional and generalized sensitivity functions with respect to x_0 and r together suggest that the beginning time interval of the solution is most sensitive to these quantities but that they are strongly correlated.

With a larger delay, $\tau = \frac{\pi}{2r} \approx 2.244$, the results given in Figure 3 reveal that many more oscillations in the solution $x(t)$ occur, although they do dampen slightly. The traditional and generalized sensitivity functions for the unknown quantities $q = (r, K, x_0, \tau)$ then indicate which parts of the oscillatory solution are most sensitive to the respective parameter q_i . Regions of decreasing GSF indicate correlation among parameters, as with the growth rate r and initial condition x_0 in Figures 1 and 2.

To illustrate the information gained from the solutions of the TSF and GSF with respect to the delay with a moderate delay $\tau = 1$, we generated simulated data with 10% error and used this to estimate the delay τ , while holding the other parameters fixed. As seen in [16], any parameter correlation issues would be irrelevant and estimates should be improved if data is concentrated in any regions of enhanced information content (regions of greatest change in GSF or TSF).

(a) model solution: with $\hat{\tau}$ using t_{unif}

(b) model solution: with $\hat{\tau}$ using t_{GSF}

Figure 4: The solutions to the delay logistic equation with estimated delay $\hat{\tau}$ from data as shown in each graph: (a) $\hat{\tau}$ with data corresponding to t_{unif} , (b) $\hat{\tau}$ with data corresponding to t_{GSF} .

The results from estimating the delay τ from 15 data points spread uniformly over the time interval $[0, 15]$ versus those for 8 out of 15 data points concentrated in the interval $[8, 11]$ are contained in Table 1. Improvement using data with enhanced information content with respect to

the delay is evident in that the estimated value for $\hat{\tau} = 1.032$ is closer to the true value $\tau^0 = 1$ than that $\hat{\tau} = 0.9561$ when uniform data is used. Additionally the resulting residual sum of squares

$$RSS = \sum_{j=1}^1 5 |y_j - f(t_j, \hat{\tau})|^2$$

is less (i.e., a better fit to data) in the estimation with data concentrated in [8, 11]. Model solutions corresponding to the estimated $\hat{\tau}$'s overlayed with the data are shown in Figure 4(a) for simulated uniform data and with data concentrated in [8, 11] in Figure 4(b).

5.2 Harmonic Oscillator

We turn finally to illustrating the use of the TSF and GSF for the Minorsky harmonic oscillators with delays as given in the Introduction. We recall that the equation with delayed damping has the form

$$\frac{d^2x(t)}{dt^2} + K \frac{dx(t-\tau)}{dt} + bx(t) = g(t), \quad (5.6)$$

while the system with delayed restoring force is given by

$$\frac{d^2x(t)}{dt^2} + K \frac{dx}{dt} + bx(t-\tau) = g(t). \quad (5.7)$$

We use traditional and generalized sensitivity functions with equations (5.6) and (5.7) and illustrate their application in determining regions of sensitivity for model parameters K, b and time delay τ . As before, we take the derivative of equation (5.6) with respect to each parameter q_i , where $q = (K, b, \tau)^T$ to obtain the TSF corresponding to that parameter q_i . First, letting $x = x_1(t)$ and $x_2(t) = \dot{x}(t)$, and rewriting equation (5.6) as a first order system we have

$$\begin{aligned} \frac{dx_1(t)}{dt} &= x_2(t) \\ \frac{dx_2(t)}{dt} &= g(t) = bx_1(t) - Kx_2(t-\tau). \end{aligned} \quad (5.8)$$

The traditional sensitivity functions are then solutions of

$$\begin{aligned} \frac{ds_1(t)}{dt} &= s_4(t) \\ \frac{ds_2(t)}{dt} &= s_5(t) \\ \frac{ds_3(t)}{dt} &= s_6(t) \\ \frac{ds_4(t)}{dt} &= -bs_1(t) - Ks_4(t-\tau) - x_2(t-\tau) \\ \frac{ds_5(t)}{dt} &= -bs_2(t) - Ks_5(t-\tau) - x_1(t) \\ \frac{ds_6(t)}{dt} &= -bs_3(t) - Ks_6(t-\tau) + K\dot{x}_2(t-\tau), \end{aligned}$$

for $s_1(t) = \frac{\partial x_1(t)}{\partial K}$, $s_2(t) = \frac{\partial x_1(t)}{\partial b}$, $s_3(t) = \frac{\partial x_1(t)}{\partial \tau}$, $s_4(t) = \frac{\partial x_2(t)}{\partial K}$, $s_5(t) = \frac{\partial x_2(t)}{\partial b}$, and $s_6(t) = \frac{\partial x_2(t)}{\partial \tau}$.

In Figure 5, the solution for the harmonic oscillator with delayed damping is shown for parameter values $K = 0.5, b = 2, g(t) \equiv 10$, and delay $\tau = 1$, along with the solutions of the traditional and

Figure 5: Depicted above are (a) the solution to the harmonic oscillator with delayed damping $K = .5, b = 2, \tau = 1$, and $g(t) = 10$, (b) the traditional sensitivity functions and (c) the generalized sensitivity functions with respect to K, b, τ .

generalized sensitivity functions with respect to $q = (K, b, \tau)^T$. The solutions of the TSFs imply that the solution is sensitive to all three parameters, with the sensitivities varying in phase with each other, and beginning when the solution itself begins oscillating. This would suggest that there may be some correlation between these three parameters, since their regions of sensitivity are identical. The solutions to the generalized sensitivity functions, however, clarify this point, and indicate that it is the parameters K and b that are correlated and the delay τ is uncorrelated with the other two over its regions of sensitivity. Therefore, if one were to estimate parameters with this model, one should not expect to estimate both K and b simultaneously, but estimating either K or b does not affect one's ability to estimate the delay τ . The solution is not sensitive to any of the parameters until the oscillations grow, indicating that data taken in the beginning time intervals should not be expected to contain much information about any of the parameters. It is not immediately obvious that two parameters K and τ appearing in the same term would be uncorrelated and therefore, both potentially are identifiable from data.

The sensitivity functions for the harmonic oscillator with delayed restoring force, equation (5.7), are arrived at in the same manner as when the delay appears in the damping term and are therefore omitted. The solution $x(t)$ and the corresponding traditional and generalized sensitivity solutions with respect to $q = (K, b, \tau)^T$ are graphed in Figure 6. The solution looks relatively monotonic, and the traditional sensitivity function solutions appear to indicate that the solution

(a) solution $x(t)$

(b) traditional sensitivity function $ts(t)$

(c) generalized sensitivity function $gs(t)$

Figure 6: Shown above are (a) the solution to the harmonic oscillator with delayed restoring force with $K = 5, b = 0.5, \tau = 1$, and $g(t) = 10$, (b) the traditional sensitivity functions and (c) the generalized sensitivity functions with respect to K, b, τ .

is disproportionately sensitive to the restoring coefficient b as compared with K and τ . However, the generalized sensitivity functions provide additional insight in that the solution appears to be sensitive to K early on and while there are regions of increased information content relative to the delay τ , that the parameter b is, not surprisingly, correlated with τ , and one should not expect to identify both simultaneously with data sampled from those intermediate and later regions.

Concluding Remarks

After giving a brief survey of previous contributions on theoretical and computational aspects of traditional sensitivity functions for delay differential equation systems, we presented a summary of new theoretical results (proofs for which are given in [27]) for differentiation of solutions with respect to parameters, initial data and delays in general nonlinear delay differential equations. These results provide a theoretical foundation for the rigorous formulation of both traditional and generalized sensitivities for delay systems. We illustrate the ideas in the context of Hutchinson's delayed logistic equation and the classical Minorsky harmonic oscillators with delayed damping or delayed restoring forces.

Acknowledgements

This research was supported in part by Grant Number NIAID R01AI071915-09 from the National Institute of Allergy and Infectious Disease, in part by the U.S. Air Force Office of Scientific Research under grant number FA9550-09-1-0226 and in part by the National Institute on Alcohol Abuse and Alcoholism under a subcontract from the Research Foundation for Mental Hygiene.

References

- [1] J. Arino, L. Wang and G. Wolkowicz, An alternative formulation for a delayed logistic equation, *J. Theo. Bio.*, **241** (2006), 109–118.
- [2] C. Baker and F. Rihan, Sensitivity analysis of parameters in modelling with delay-differential equations, *MCCM Tec. Rep.*, **349** (1999), Manchester, ISSN 1360-1725.
- [3] H.T. Banks, *Modeling and Control in the Biomedical Sciences*, Lecture Notes in Biomath., Vol. 6, Springer-Verlag, Berlin, Heidelberg, New York, 1975.
- [4] H.T. Banks, Delay systems in biological models: approximation techniques, *Nonlinear Systems and Applications* (V. Lakshmikantham, ed.), Academic Press, New York, 1977, 21–38.
- [5] H. T. Banks, Approximation of nonlinear functional differential equation control systems, *J. Optimiz. Theory Appl.*, **29** (1979), 383–408.
- [6] H. T. Banks, Identification of nonlinear delay systems using spline methods, in: V. Lakshmikantham (Ed.), *Nonlinear Phenomena in Mathematical Sciences*, Academic Press, New York, NY, 1982, 47–55.
- [7] H. T. Banks, *A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering*, CRC Press/Taylor and Frances Publishing, Accepted, January 15, 2012. (258 pages)

- [8] H.T. Banks and J.A. Burns, An abstract framework for approximate solutions to optimal control problems governed by hereditary systems, *International Conference on Differential Equations* (H. Antosiewicz, Ed.), Academic Press (1975), 10–25.
- [9] H.T. Banks and J.A. Burns, Hereditary control problems: numerical methods based on averaging approximations, *SIAM J. Control & Opt.*, **16** (1978), 169–208.
- [10] H.T. Banks, Identification of nonlinear delay systems using spline methods. In *Nonlinear Phenomena in Mathematical Sciences*, V. Lakshmikanitham, Ed. Academic Press, Inc., New York, NY, 1982, pp. 47–45.
- [11] H.T. Banks and D. M. Bortz, A parameter sensitivity methodology in the context of HIV delay equation models, *Journal of Mathematical Biology*, **50(6)** (2005), 607–625.
- [12] H.T. Banks, J.A. Burns and E.M. Cliff, Parameter estimation and identification for systems with delays, *SIAM J. Control and Optimization*, **19** (1981), 791–828.
- [13] H.T. Banks, M. Davidian, J.R. Samuels, Jr. and Karyn L. Sutton, An inverse problem statistical methodology summary, CRSC-TR08-01, January, 2008; Chapter 11 in *Mathematical and Statistical Estimation Approaches in Epidemiology*, (edited by Gerardo Chowell, et al.), Springer, Berlin Heidelberg New York, 2009, pp. 249–302.
- [14] H.T. Banks, J.E. Banks and S.L. Joyner, Estimation in time-delay modeling of insecticide-induced mortality, CRSC-TR08-15, October, 2008; *J. Inverse and Ill-posed Problems*, **17** (2009), 101–125.
- [15] H. T. Banks, S. Dediu and S. L. Ernstberger, Sensitivity functions and their uses in inverse problems, CRSC-TR07-12, July, 2007; *Journal of Inverse and Ill-Posed Problems*, **15** (2007), 683–708.
- [16] H.T. Banks, S. Dediu, S. Ernstberger and F. Kappel, Generalized sensitivities and optimal experimental design, CRSC-TR08-12 (Revised), November, 2009; *J. Inverse and Ill-posed Problems*, **18** (2010), 25–83.
- [17] H.T. Banks, S. Dediu and H.K.Nguyen, Time delay systems with distribution dependent dynamics, CRSC-TR06-15, May, 2006; *IFAC Annual Reviews in Control*, **31** (2007), 17–26.
- [18] H.T. Banks, S. Dediu and H. K. Nguyen, Sensitivity of dynamical systems to parameters in a convex subset of a topological vector space, CRSC-TR06-25, November, 2006; *Mathematical Biosciences and Engineering*, **4** (2007), 403–430.
- [19] H.T. Banks, K. Holm and F. Kappel, Comparison of optimal design methods in inverse problems, CRSC-TR10-11, July, 2010; *Inverse Problems*, **27** (2011) 075002(31pp).
- [20] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, *J. Differential Equations*, **34** (1979), 496–522.
- [21] H.T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhausen, Bosten, 1989.
- [22] H.T. Banks and P.K. Daniel Lamm, Estimation of delays and other parameters in nonlinear functional differential equations, *SIAM J. Control and Optimization*, **21** (1983), 895–915.

- [23] H.T. Banks and J.M. Mahaffy, Global asymptotic stability of certain models for protein synthesis and repression, *Quart. Applied Math.*, **36** (1978), 209–221.
- [24] H.T. Banks and J.M. Mahaffy, Stability of cyclic gene models for systems involving repression, *J. Theoretical Biology*, **74** (1978), 323–334.
- [25] H.T. Banks and H. Nguyen, Sensitivity of dynamical systems to Banach space parameters, CRSC-TR05-13, February, 2005; *J. Math. Analysis and Applications*, **323** (2006), 146–161.
- [26] H.T. Banks, K. Rehm and K. Sutton, Inverse problems for nonlinear delay systems, CRSC-TR10-17, N.C. State University, November, 2010; *Methods and Applications of Analysis*, **17** (2010), 331–356.
- [27] H.T. Banks, D. Robbins and K. Sutton, Theoretical foundations for traditional and generalized sensitivity functions for nonlinear delay differential equations, to appear.
- [28] H.T. Banks and I.G. Rosen, Spline approximations for linear nonautonomous delay systems, ICASE Rep. No. 81-33, NASA Langley Res. Center, Oct., 1981; *J. Math. Anal. Appl.*, **96** (1983), 226–268.
- [29] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Noordhoff, Leyden, 1976.
- [30] R. Bellman and K. L. Cooke, *Differential-Difference Equations*, Vol. **6**, Mathematics in Science and Engineering, Academic Press, New York, NY, 1963.
- [31] D. Brewer, The differentiability with respect to a parameter of the solution of a linear abstract Cauchy problem, *SIAM J. Math. Anal.*, **13** (1982), 607–620.
- [32] D. Brewer, J.A. Burns, and E. M. Cliff, Parameter Identification for An Abstract Cauchy Problem by Quazilinearization, ICASE Rep. No. 89-75, NASA Langley Res. Center, Oct., 1989.
- [33] M.D. Buhmann and A. Iserles, On the dynamics of a discretized neutral equation, *IMA J. of Numerical Analysis*, **12** (1992), 339–363.
- [34] J.A. Burns, E.M. Cliff and S.E. Doughty, Sensitivity analysis and parameter estimation for a model of Chlamydia Trachomatis infection, *J. Inverse Ill-Posed Problems*, **15** (2007), 19–32.
- [35] S.N. Busenberg and K.L. Cooke, eds., *Differential Equations and Applications in Ecology, Epidemics, and Population Problems*, Academic Press, New York, 1981.
- [36] V. Capasso, E. Grosso and S.L. Paveri-Fontana, eds., *Mathematics in Biology and Medicine*, Lecture Notes in Biomath., Vol. 57, Springer-Verlag, Berlin, Heidelberg, New York, 1985.
- [37] J. Caperon, Time lag in population growth response of *Isochrysis Galbana* to a variable nitrate environment, *Ecology*, **50** (1969), 188–192.
- [38] S. Choi and N. Koo, Oscillation theory for delay and neutral differential equations, *Tends in Mathematics*, **2** (1999), 170–176.
- [39] K.L. Cooke, Functional differential equations: Some models and perturbation problems, in *Differential Equations and Dynamical Systems* (J.K. Hale and J.P. LaSalle, eds.), Academic Press, New York, 1967, 167–183.

- [40] J.M. Cushing, *Integro-differential Equations and Delay Models in Population Dynamics*, Lec. Notes in Biomath., Vol. 20, Springer-Verlag, Berlin and New York, 1977.
- [41] C. Elton, *Voles, Mice, and Lemmings*, Clarendon Press, London, 1942.
- [42] P.L. Errington, Predation and vertebrate populations, *Quarterly Review of Biology*, **21** (1946), 144–177, 221–245.
- [43] U. Forys and A. Marciniak-Czochra, Delay logistic equation with diffusion, *Proc 8th Nat. Conf. Mathematics Applied to Biology and Medicine*, Lajs (2002), Warsaw, Poland, 37–42
- [44] U. Forys and A. Marciniak-Czochra, Logistic equations in tumor growth modelling, *Int.J. Appl. Math. Comput. Sci.*, **13** (2003), 317–325.
- [45] J.S. Gibson and Clark, Sensitivity analysis for a class of evolution equations, *J. Mathematical Analysis and Applications*, **58** (1977), 22–31.
- [46] L. Glass and M. C. Mackey, Pathological conditions resulting from instabilities in physiological control systems, *Ann. N. Y. Acad. Sci.* **316** (1979), 214–235.
- [47] K. P. Hadeler, Delay equations in biology, in *Functional differential equations and approximation of fixed points*, **730** Springer-Verlag Berlin, 1978, 136–156.
- [48] J. K. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, 1977.
- [49] F.C. Hoppensteadt, ed., *Mathematical Aspects of Physiology*, Lectures in Applied Math, Vol. 19, American Mathematical Society, Providence, 1981.
- [50] G.E. Hutchinson, Circular causal systems in ecology, *Ann. N.Y. Acad. Sci.*, **50** (1948) 221–246.
- [51] G. E. Hutchinson, *An Introduction to Population Ecology*, Yale University, New Haven, 1978.
- [52] F. Kappel, An approximation scheme for delay equations, in V. Lakshmikantham (Ed.), *Non-linear Phenomena in Mathematical Sciences*, Academic Press, New York, NY, 1982, 585–595.
- [53] F. Kappel, Generalized sensitivity analysis in a delay system, *Proc. Appl. Math. Mech.*, **7** (2007), 1061001–1061002.
- [54] F. Kappel and W. Schappacher, Autonomous nonlinear functional differential equations and averaging approximations, *J. Nonlinear Analysis*, **2** (1978), 391–422.
- [55] Y. Kuang, *Delay Differential Equations: with Applications in Population Dynamics*, Academic Press, Inc., San Diego, 1993.
- [56] N. MacDonald, Time lag in a model of a biochemical reaction sequence with end-product inhibition, *J. Theor. Biol.*, **67** (1977), 727–734.
- [57] M. Martelli, K.L. Cooke, E. Cumberbatch, B. Tang and H. Thieme, eds., *Differential Equations and Applications to Biology and to Industry*, World Scientific, Singapore. 1996.
- [58] J.A.J. Metz and O. Diekmann, eds., *The Dynamics of Physiologically Structured Populations*, Lecture Notes in Biomath., Vol. 68, Springer-Verlag, Berlin, Heidelberg, New York, 1986.
- [59] N. Minorsky, Self-excited oscillations in dynamical systems possessing retarded actions, *J. Applied Mechanics*, **9** (1942), A65–A71.

- [60] N. Minorsky, On non-linear phenomenon of self-rolling, *Proc. National Academy of Sciences*, **31** (1945), 346–349.
- [61] N. Minorsky, *Nonlinear Oscillations*, Van Nostrand, New York, 1962.
- [62] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
- [63] D.M. Pratt, Analysis of population development in *Daphnia* at different temperatures, *Biology Bulletin*, **22** (1943), 345–365.
- [64] D. Robbins, *Sensitivity Functions for Delay Differential Equation Models*, Ph D Dissertation, North Carolina State University, Raleigh, September, 2011.
- [65] K. Schmitt, ed., *Delay and Functional Differential Equations and Their Applications*, Academic Press, New York, 1972.
- [66] R. Schuster and H. Schuster, Reconstruction models for the Ehrlich Ascites Tumor of the mouse, in *Mathematical Population Dynamics, Vol. 2* (edited by O. Arino, D Axelrod, M. Kimmel), Wuertz, Winnipeg, 1995, 335–348.
- [67] L.F. Shampine and S. Thompson, *Delay Differential Equations With dde23*, (2000), SMU/RU.
- [68] F. R. Sharpe and A. J. Lotka, Contribution to the analysis of malaria epidemiology IV: Incubation lag, supplement to *Amer. J. Hygiene* **3** (1923), 96–112. (reprinted in [71] Scudo and Ziegler (1978) -).
- [69] M.H. Schultz, *Spline Analysis*, Prentice- Hall, Englewood Cliffs, N.J., 1973.
- [70] M.H. Schultz and R.S. Varga, L-splines, *Num. Math.*, **10** (1967), 345–369.
- [71] F. M. Scudo and J. R. Ziegler, *The Golden Age of Theoretical Ecology*, Springer, Berlin 1978.
- [72] K. Thomseth and C. Cobelli, Generalized sensitivity functions in physiological system identification, *Ann. Biomed. Engr.*, **27** (1999), 606–616.
- [73] E.M. Wright, A non-linear difference-differential equation, *J. Reine Angew. Math*, **494** (1955), 66–87.