Sensitivity of Dynamical Systems to Parameters in a Convex Subset of a Topological Vector Space

H. T. Banks[a,*], Sava Dediu[a,b,†], and Hoan K. Nguyen[a,‡]

[a]Center for Research in Scientific Computation

Box 8205

N. C. State University

Raleigh, NC 27695-8205

and

[b]Statistical and Applied Mathematical Sciences Institute

19 T. W. Alexander Dr

P. O. Box 14006

Research Triangle Park, NC 27709-4006

November 15, 2006

Abstract

We develop a theory for sensitivity with respect to parameters in a convex subset of a topological vector space of dynamical systems in a Banach space. Specific motivating examples for probability measure dependent differential, partial differential and delay differential equations are given. Schemes that approximate the measures in the Prohorov sense are illustrated with numerical simulations for distributed delay differential equations.

1   Introduction

In this paper we develop sensitivity equations for a general nonlinear dynamical system in a Banach space depending on parameters in a convex subset of a topological vector space

---

[*]email: htbanks@ncsu.edu

[†]email: sdediu@ncsu.edu

[‡]email: hknguyen@ncsu.edu

which is not necessarily a normed linear space. More specifically, we study the sensitivity equation of the following ordinary differential equation

$$\dot{x}(t) = f(t, x(t), \mu) \qquad \text{if } t \geq t_0,$$
$$x(t_0) = x_0, \tag{1}$$

whose solution x is assumed to exist and belong to a complex Banach space $(X, |\cdot|_X)$ for any choice of the parameter μ in a convex subset $\mathcal{M}$ of a topological vector space $\mathcal{X}$. Our investigations are motivated by the fact that in many applications of practical interest, the parameter space or set is not a Banach space or even a linear space and the sensitivity analysis developed in [19] is therefore not applicable.

As we shall see below, our motivating example involves the set $\mathcal{P}(S)$ (or a convex subset of $\mathcal{P}(S)$), the general metric space of probability measures defined on a set S, taken with the Prohorov metric (discussed in detail below), when $\mathcal{P}(S)$ is viewed as a subset of the dual $\mathcal{X} = (C_B^*(S), \mathcal{T}_{w^*})$ of the bounded continuous functions $C_B(S)$, with $\mathcal{X} = C_B^*(S)$ taken with the weak* topology.

The topological vector space structure of $\mathcal{X}$ and the underlying convex subset $\mathcal{M} \subset \mathcal{X}$ do not permit the usual framework where we can define and use a Frechet derivative of f with respect to μ. However, the convexity assumption for $\mathcal{M}$ allows us to perform a sensitivity analysis of (1) by means of directional derivatives of f with respect to μ, when μ stays within the convex set $\mathcal{M}$. Given two points μ, ν in $\mathcal{M}$, we define the derivative $\delta f(t, x, \mu; \nu - \mu)$ of f at μ *in the direction* $\nu - \mu$ to be the value of the limit

$$\lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \frac{f(t, x, \mu + \epsilon(\nu - \mu)) - f(t, x, \mu)}{\epsilon} = \delta f(t, x, \mu; \nu - \mu), \tag{2}$$

provided this limit exists in X. Our goal in this paper is to show that under suitable conditions imposed on f, *the sensitivity of the solution x with respect to μ*, defined as the directional derivative

$$y(t) = \delta x(t, \mu; \nu - \mu) = \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \frac{x(t, \mu + \epsilon(\nu - \mu)) - x(t, \mu)}{\epsilon}, \tag{3}$$

exists in X and satisfies the *sensitivity equation*

$$\dot{y}(t) = f_x(t, x(t, \mu), \mu)y(t) + \delta f(t, x(t, \mu), \mu; \nu - \mu) \qquad \text{if } t \geq t_0,$$
$$y(t_0) = 0. \tag{4}$$

We note the different nature of the derivatives of f appearing above. Here f_x is the Frechet derivative of f with respect to x, whereas $\delta f(t, x, \mu; \nu - \mu)$ is the directional derivative of f at μ in the direction $\nu - \mu$.

Motivation for investigating sensitivity is widespread and there is a substantial literature on the subject. The large literature includes a number of books devoted to both elementary

and advanced aspects of sensitivity [34, 36, 38, 40, 41, 50, 61]. Equations for the sensitivity of a system with respect to finite dimensional vector parameters are used in standard methodology for optimization and inverse problems including least squares, maximum likelihood, computation of standard errors in statistics, etc., [35], as well as in model discrimination/model selection related quantities (dispersion matrix, Fisher information matrix, etc., [29, 44]). Applications of sensitivity methods for finite dimensional vector parameters are widespread and can be found in many applications including biology [24] and physiology [22], mechanics [1, 41], and control theory [61]. More recently, investigators' attention has turned to more complex formulations for sensitivity of infinite dimensional systems with function space parameters in problems involving shape sensitivities or sensitivities with discontinuous coefficients (for interesting and well motivated pioneering results in this area, we recommend that the reader see [25, 26, 30, 31, 32, 33] and [54] and the references therein). However, sensitivity for the systems we investigate here of the form

$$\dot{x}(t) = F(t, x(t), P), \tag{5}$$

where P is a probability distribution or measure, represents another class of problems where the "parameter" sets are infinite dimensional and presents a new set of both theoretical and computational challenges. Again the applications motivating such systems are diverse and growing with the increased scientific interest in incorporating uncertainty into models and systems. The need to employ dynamics with probabilistic structures has recently received increased emphasis. In particular, systems with probability measures or distributions embedded in the dynamics (problems involving *aggregate dynamics* as discussed in [8]) have become important in applications in biology [6, 7, 8], electromagnetics [11, 12] and hysteretic [15, 16, 28, 43, 45] and polymeric [17, 18, 20] materials. While the interest in sensitivity of systems such as (5) is recent, we point out that general interest in such systems is rather longstanding. We illustrate with a simple example from control theory arising in *relaxed* or *chattering* control problems [46, 47, 51, 58, 59, 60] wherein the controls are probability measures. Indeed, such systems date back to the seminal work of L.C. Young on *generalized curves* in the calculus of variations [62, 63].

Briefly, early applications in what is now called classical optimal control theory led to the need for controls that switched frequently or "chattered" between constant (in time) values in a control restraint set U. To be more precise, one has a dynamical system of the form

$$\frac{dx}{dt}(t) = f(t, x(t), u(t)), \quad 0 \le t \le T, \tag{6}$$

and one wishes to "control" the system through the input or control function u which takes values in the set U. Here the requirement that $u(t) \in U, 0 \le t \le T$, is dictated by physics or the engineering aspects of the problem. In the so-called "chattering controls" that were found to be "near optimal" (for certain given criteria), the controls were piecewise constant taking values u_j on intervals $[t_j, t_{j+1})$ comprising a partition of $[0, T]$. Thus the near optimal controls had the form

$$u(t; v) = \chi_{\{u_j\}}(v)\chi_{[t_j, t_{j+1})}(t) \text{ for } v \in U, \ 0 \le t \le T. \tag{7}$$

In differential form (with respect to the control set) this becomes

$$du(t; v) = \delta_{\{u_j\}}(v)dv\chi_{[t_j, t_{j+1})}(t). \tag{8}$$

This in turn leads to the near optimal form of (6) being given by

$$\frac{dx}{dt}(t) = \int_U f(t, x(t), v)\delta_{\{u_j\}}(v)dv\chi_{[t_j, t_{j+1})}(t) = f(t, x(t), u_j)\chi_{[t_j, t_{j+1})}(t). \tag{9}$$

Furthermore, it was found that by increasing the number of switches through refinement of the time partition (so that the control switches or "chatters" increasingly often), one does increasingly better in optimizing the cost criterion. In the limit one obtains an optimal control that switches infinitely often between certain values in U. Such a control "function" is not a function in the usual sense but it can be realized by a "relaxed control" or U-valued time dependent measure $\mu(t; \cdot)$ defined on the Borel subsets $\mathcal{B}$ of U. The corresponding limiting control system becomes

$$\frac{dx}{dt}(t) = \int_U f(t, x(t), v)d\mu(t; v) = F(t, x(t), \mu(t)). \tag{10}$$

Here the time dependent measure μ is the $weak^* - limit$ (on $C_B(U)$, the bounded continuous functions $v \rightarrow f(t, x, v)$) as the partition mesh in (9) approaches zero. By formulating the optimization problem over the *relaxed control system* (10), one obtains not only *closure* but also *convexity* in the resulting control problem. In this relaxed control setting it is then possible to obtain well-posedness (existence, uniqueness, appropriate continuous dependence, etc.) for the optimal control problem and to argue that the resulting optimal limit or relaxed control can be well approximated by sequences of ordinary control functions of the form given in (7). While much theoretical work (existence, uniqueness and some approximation) was carried in the period $1960 - 1975$, little to no effort was made on a sensitivity methodology for such problems. While the by-now classical references (in particular, the book [60] by Warga presents a beautiful summary) are difficult to read, such an undertaking is well-worth the effort for the wealth of ideas found in this area of control theory. We note that the weak* topology on the dual $C_B^*(U)$ was quite naturally used in the development of relaxed control theory and, as we shall explain below, this is precisely the topology induced by the Prohorov metric (although convergence of measures in this topology is often called "weak convergence" by probabilists in something of a misnomer from a functional analysis point of view). We also remark here that more recent [15, 16] formulations for *materials with memories* in a *Preisach measures* setting also lead to systems of the form (10) and provide yet another motivation for the sensitivity efforts pursued in this paper. Before turning to our new theoretical and computational findings for these systems, in the next section we offer several current applications where measure or probability distribution dependent dynamical systems are playing a major role.

2 Motivating Examples with Measure/Probability Density Dependent Dynamics

Our first example from recent investigations [11, 12] in which the systems of the type under investigation here arise naturally entails Maxwell's equations for electromagnetic waves impinging on a heterogeneous dielectric material such as biotissue. If we consider a domain $\mathcal{D} = \Omega_0 \cup \Omega$ where the dielectric material is Ω and the ambient Ω_0 is treated as a vacuum, then Maxwell's equations govern the electric field $\mathbf{E}$ and the magnetic field $\mathbf{H}$ in $\mathcal{D}$ and are given by the system

$$
\begin{cases}
\text{(i)} & \dfrac{\partial \mathbf{D}}{\partial t} + \mathbf{J} - \nabla \times \mathbf{H} = 0, \text{ in } (0, T) \times \mathcal{D}, \\[2mm]
\text{(ii)} & \dfrac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0, \text{ in } (0, T) \times \mathcal{D}, \\[2mm]
\text{(iii)} & \nabla \cdot \mathbf{D} = \rho, \text{ in } (0, T) \times \mathcal{D}, \\[2mm]
\text{(iv)} & \nabla \cdot \mathbf{B} = 0, \text{ in } (0, T) \times \mathcal{D}, \\[2mm]
\text{(v)} & \mathbf{E} \times \mathbf{n} = 0, \text{ on } (0, T) \times \partial\mathcal{D}, \\[2mm]
\text{(vi)} & \mathbf{E}(0, \mathbf{x}) = 0, \quad \mathbf{H}(0, \mathbf{x}) = 0, \text{ in } \mathcal{D}.
\end{cases}
\tag{11}
$$

Here ρ is the charge density in Ω and, as usual, the current $\mathbf{J}$ is composed of a source current $\mathbf{J}_s$ and a material induced conductive current $\mathbf{J}_c$ in Ω. Within the domain $\mathcal{D}$ we need *constitutive relations* that relate the flux densities $\mathbf{D}, \mathbf{B}$ and the conductive current $\mathbf{J}_c$ to the electric and magnetic fields $\mathbf{E}$ and $\mathbf{H}$, respectively. These are given by in Ω by

$$
\begin{cases}
\text{(i)} & \mathbf{D} = \epsilon_0 \epsilon_r \mathbf{E} + \mathbf{P_m}, \\[2mm]
\text{(ii)} & \mathbf{B} = \mu_0 \mathbf{H}, \\[2mm]
\text{(iii)} & \mathbf{J}_c = \sigma \mathbf{E},
\end{cases}
\tag{12}
$$

where $\epsilon_r = (1 + \chi)$ is the relative permittivity of the dielectric medium and χ accounts for the instantaneous polarization due to the interface between Ω_0 and Ω. (Of course in Ω_0 we have $\mathbf{J}_c = 0, \mathbf{P_m} = 0$ and $\chi = 0$ so $\epsilon_r = 1$.) The dielectric material polarization $\mathbf{P_m}$ generally has convolution form

$$
\mathbf{P_m}(t, \mathbf{x}) = g * \mathbf{E}(t, \mathbf{x}) = \int_0^t g(t - s, \mathbf{x}) \mathbf{E}(s, \mathbf{x}) ds,
\tag{13}
$$

where g is the general dielectric response function (DRF). In every practical example (Debye, Lorentz, etc.) DRFs are parameter dependent as well as time (and possibly space) dependent; we represent this as $g = g(t, \mathbf{x}; \nu)$, where typically $\nu = (\epsilon_\infty, \epsilon_s, \tau)$ contains parameters such as the high frequency limit dielectric permittivity ϵ_∞, the static permittivity ϵ_s, and relaxation time τ. Examples of often-used DRFs (see [9]) are the Debye in a material region Ω defined in the time domain by

$$
g(t, \mathbf{x}) = \epsilon_0 (\epsilon_s - \epsilon_\infty)/\tau \ e^{-t/\tau},
$$

5

and the Lorentz given by
$$g(t, \mathbf{x}) = \epsilon_0 \omega_p^2 / \nu_0 e^{-t/2\tau} sin(\nu_0 t).$$

The macroscopic polarization model (13) can be derived from microscopic dipole, electron cloud, etc., formulations by passing to a limit over the molecular population. However, such derivations tacitly assume that one has similar individual (molecular, dipole, etc.) parameters; that is, all dipoles, molecules, "electron clouds", etc., have the same relaxation parameters, plasma frequencies, etc. Historically, such models based on molecular level homogeneity throughout the material have often not performed well when trying to compare models with experimental data. Indeed, in 1907 Von Schweidler [27, 56] observed the need to assume *multiple relaxation times* when considering experimental data and in 1913 Wagner [27, 57] proposed continuous distributions of relaxation times. There are now *incontrovertible* experimentally based arguments for distributions of relaxation parameters in mechanisms for heterogeneous materials (see [37]). Moreover, there is compelling evidence of the presence of *multiple mechanisms* in complex materials such as tissue and modern polymeric composites. These multiple mechanisms may involve interfacial polarization, dipolar orientation, ionic diffusion and may often require a selection of several types of distributional representations from examples such as the Debye and Lorentz models. To allow for a distribution P_1 of parameters ν over some admissible set $\mathcal{N}$, it is useful to generalize the polarization law (13) to
$$\mathbf{P_m}(t, \mathbf{x}; P_1) = (h(P_1) * \mathbf{E})(t, \mathbf{x}) = \int_0^t \int_{\mathcal{N}} g(t - s, \mathbf{x}; \nu) \mathbf{E}(s, \mathbf{x}) dP_1(\nu) ds, \qquad (14)$$
where
$$h(t, \mathbf{x}; P_1) = \int_{\mathcal{N}} g(t, \mathbf{x}; \nu) dP_1(\nu). \qquad (15)$$

We expect to chose P_1 from (or from a subset of) the space $\mathcal{P}(\mathcal{N})$ of all probability measures P_1 on $\mathcal{N}$.

To allow for dielectric materials with multiple mechanisms or multiple DRFs (i.e., heterogeneous molecular structures), one may further generalize (14) by considering a family $\mathcal{G}$ of possible DRFs and distributions P_2 over this family. This leads to the polarization constitutive relationship

$$\mathbf{P_m}(t, \mathbf{x}; P_2, P_1) = \int_0^t \int_{\mathcal{G}} \int_{\mathcal{N}} g(t - s, \mathbf{x}; \nu) \mathbf{E}(s, \mathbf{x}) dP_1(\nu) dP_2(g) ds$$
$$= \int_0^t K(t - s, \mathbf{x}; P_2, P_1) \mathbf{E}(s, \mathbf{x}) ds, \qquad (16)$$

where for $P_1 \in \mathcal{P}(\mathcal{N})$ and $P_2 \in \mathcal{P}(\mathcal{G})$, K is defined by

$$K(t - s, \mathbf{x}; P_2, P_1) = \int_{\mathcal{G}} \int_{\mathcal{N}} g(t - s, \mathbf{x}; \nu) dP_1(\nu) dP_2(g). \qquad (17)$$

When we use (16)–(17) in the Maxwell system (11)–(12), we are led to a system of partial differential equations where lower order terms (in time) depend on probability measures.

These resulting systems have the form of a partial differential equation which can be written as an abstract ordinary differential equation in a Banach or Hilbert space

$$\dot{x}(t) \;=\; F(t, x(t), P_1, P_2),$$

where P_1, P_2 are probability distributions or measures.

A second class of models [5] depending explicitly on distributions to incorporate uncertainty in the system dynamics involves the use of shrimp grown in production "raceways" (essentially large growth chambers wherein environmental factors such as temperature, oxygen, nutrient levels, etc., can be carefully controlled) artificially infected to efficiently produce large quantities of an associated vaccine. In such systems one recruits the biochemical machinery in an existing biomass for the production of a vaccine or antibody by infection using a virus carrying a passenger gene for the desired antibody response.

In the example developed in [5] one begins with a population of healthy shrimp and infects them with a recombinant viral vector (e.g., recombinant Taura Syndrome virus or rTSV in the example developed in [5]) expressing a foreign antigen, resulting in vaccine production in live infected shrimp. For simplicity the amount of vaccine produced can be assumed equal to the total infected biomass so that vaccine production will essentially follow the course of the viral dynamics in the shrimp. This requires modeling the dynamics of shrimp at the population level. Since shrimp have size dependent characteristics as well as structure-dependent responses to external environment, an appropriate model might be based on the classical McKendrick-von-Foerster/Sinko-Streifer size-structured population equations [42, 48] with mass as the structure variable, i.e., one equates the size variable with the mass in this model.

Experimental results suggest that mortality rates in acutely infected shrimp depend on the length of time that the shrimp remain acute. Moreover individuals in a latent phase have varying residency times before they progress into the acute phase. To account for the different residency periods of individual shrimp one can record the variable residency times in the different stages by introducing a new variable which one calls the *class age* of an individual. The class age of an individual in a given stage represents the length of time that the individual spends in that stage and serves as a surrogate for time delays.

In the vaccine production stage the shrimp can be infected by distributing chopped dead shrimp infected with a recombinant virus evenly throughout the raceway and it might be reasonably assumed that all the shrimp have an equal chance of becoming infected by eating the infected biomass. The expected time interval for infection is 7 to 10 days. During this time interval almost no shrimp progress into the chronic state. Therefore it is reasonable to consider the following three compartment states: susceptible (S), latently infected (L) and acutely infected (A) in a model (as depicted schematically in Figure 1) where it is assumed that shrimp will become instantly infected (i.e., progress into latent state) as soon as they ingest some of the infected biomass. As we have noted earlier, however, experimental observations suggest that there exists a temporal delay between the initial latent infection and initial acute infection [39]. Moreover, it is biologically unrealistic to expect all members of the shrimp population to progress into the acute phase at a fixed number of days after
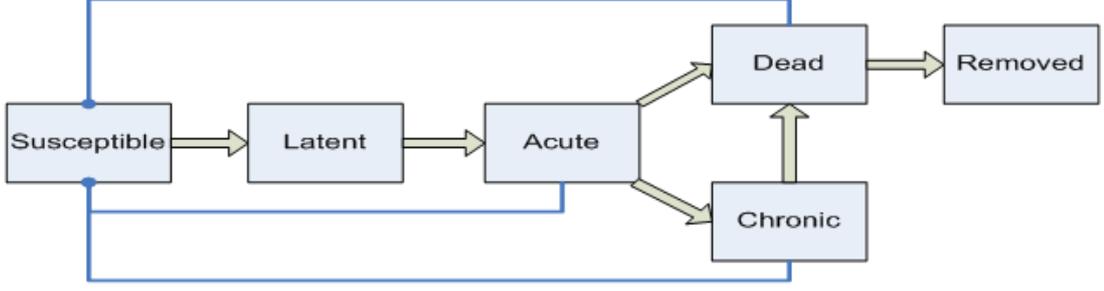
Figure 1: Stages in TSV-infected shrimp.

initial latent infection. In addition the shrimp in the acute phase have varying mortality rates because of the different times that they progress into the acute phase and also due to the differences in genetic make-up of the host. It is difficult to account for the class age history (i.e., the length of time that shrimp spend in a state) of shrimp in a particular (latent or acute) state using a system of delay PDE's with only size as the structure variable. So in order to model variable residency times one may keep track of the class age and the size of shrimp by incorporating both size structure and class age structure into the latent and acute states.

It is expected that there is a positive probability that shrimp can stay in each the latent and acute state for more than 7 to 10 days. Thus one can assume that the class age interval for both states is the same as the time interval T_V that we consider in our model.

The vaccine production model based on the above discussions, the stages as depicted in Figure 1 and as developed in [5], is given by

$$\partial_t S(x,t) + \partial_x(g^S(x)S(x,t)) + m^S(x)S(x,t) = -\lambda S(x,t),$$

$$\partial_t L(x,t,\theta) + \partial_x(g^L(x)L(x,t,\theta)) + \partial_\theta L(x,t,\theta) + m^L(x)L(x,t,\theta) = -p_L(\theta)L(x,t,\theta),$$

$$\partial_t A(x,t,\theta) + \partial_\theta A(x,t,\theta) + m^A(\theta)A(x,t,\theta) = 0,$$

$$L(x,t,0) = \lambda S(x,t),$$

$$A(x,t,0) = \int_0^t L(x,t,\theta)dP_L(\theta),$$

$$S(0,t) = 0, \quad L(0,t,\theta) = 0, \quad A(0,t,\theta) = 0,$$

$$S(x,0) = S^0(x), \quad L(x,0,\theta) = 0, \quad A(x,0,\theta) = 0,$$

where $(x,t,\theta) \in [x_{\min}, x_{\max}] \times [0, T_V] \times [0, T_V]$. In the above $S(x,t)$ denotes the density of susceptible individuals having mass x at time t and $\partial_t = \frac{\partial}{\partial t}$. The function $L(x,t,\theta)$ denotes the density of latent individuals having mass x at time t that have spent θ days in the
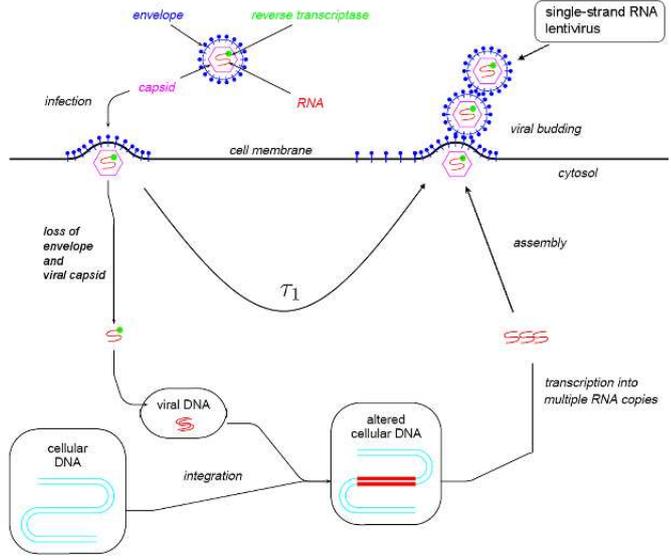
8

Figure 2: HIV infection pathway in acutely infected cells.

latent state, whereas the function $A(x, t, \theta)$ denotes the density of acute individuals having mass x at time t that have spent θ days in the acute state. The quantity $g^S(x)$ denotes the growth rate of individuals in the susceptible state and $g^L(x)$ denotes the growth rate of individuals in the latent state. The function $m^S(x)$ denotes the mortality rate of individuals in the susceptible state, the function $m^L(x)$ denotes the mortality rate of individuals in the latent state, and $m^A(\theta)$ denotes the mortality rate of the shrimp that spend θ days in the acute state. The latent to acute probability density rate function $p_L(\theta) = P'_L(\theta)$ defined for $\theta \in [0, T_V]$ denotes the rate at which the shrimp in the latent state that have spent θ days in the latent state become acutely infected, while the quantity λ denotes the infection rate due to ingestion of chopped infected shrimp. Finally $S^0(x)$ denotes the initial population density of susceptible shrimp produced from the biomass production model.

We note that this is again a probability distribution (P_L) dependent dynamical system (in this case a complicated system of partial differential equations) for which the distribution P_L must be estimated in some type of inverse problem.

Our next example is one developed and investigated in [6, 7] for models of progression of Human Immunodeficiency Virus (HIV) at the cellular level in mice. The resulting system is typical of delay systems that arise in biochemical pathways and cellular level kinetics of drug metabolism as well as other synthesis models. The HIV system based on the schematic

9

for activity within an acutely infected cell as depicted in Figure 2 has the form

$$
\begin{aligned}
\dot{V}(t) &= -cV(t) + n_A \int_{-\infty}^{0} A(t+\theta)dP_1(\theta) + n_C C(t) - n_I V(t)T(t) \\
\dot{A}(t) &= (r_v - \delta_A - \delta X(t))A(t) + n_I V(t)T(t) - \gamma \int_{-\infty}^{0} A(t+\theta)dP_2(\theta) \\
\dot{C}(t) &= (r_v - \delta_A - \delta X(t))C(t) + \gamma \int_{-\infty}^{0} A(t+\theta)dP_2(\theta) \\
\dot{T}(t) &= (r_v - \delta_A - \delta X(t) - n_I V(t))T(t) + S,
\end{aligned}
\tag{18}
$$

which is a vector system of integro-differential equations. Special cases of such systems include those in which the probability measures are defined on some finite interval $Q = [-r, 0]$ of possible delay values θ. In this model one has compartments $T, A, C,$ and V for *in vitro* blood level counts in mice of target $(CD4+)$ cells, acutely infected cells, chronically infected cells and active viral particles, respectively. Free virus V infects target cells T, transforming them into acutely infected cells A which at some time later become chronically infected cells C. The basic pathway for infection and production of virus for acutely infected cells is schematically depicted in Figure 2. For models in which the individual kinetics for loss of envelope and capsid, integration, transcription, and assembly are not detailed, it is necessary (see [7]) to include a delay τ_1 from the time of infection of a target cell T until it first produces free virus V. There is also some delay τ_2 before an acutely infected cell A as depicted in Figure 2 becomes a chronically infected cell C. However, these delays vary over the population of cells and hence it is the *distributions* P_1, P_2 of these delays, not a specific value of the delays that are of interest in the modeling. The variables C and V then become the expected values of the chronic cell levels and the viral loads, respectively (see [7] for a detailed derivation). This model was successfully used (again see [7]) to describe *in vitro* mice data from Dr. Michael Emerman's lab at Fred Hutchinson Cancer Research Center. Using inverse problem methodology and statistical analysis it was shown that improvement of fit to data by inclusion of the delays τ_1 or P_1 is statistically significant while inclusion of delays τ_2 (or distributions P_2) is less important. Indeed, it was found that the experimental data could not be properly fit with an ODE version (i.e., with the delay distributions omitted) of the model.

In a final example we introduce briefly a model for sublethal damage in populations of insects (pea aphids in this case) being exposed to pesticides. Background for such models can be found in [3] where both Leslie matrix models and size-structured partial differential equation models are employed. Here we discuss a different approach involving models we are currently investigating; this formulation focuses on uncertainty with respect to the delay after exposure before death. While the exact mechanisms for death are still being investigated by entomologists, there is sufficient understanding to attempt to develop and test models against experimental data. Insecticides generally affect neonates and adults somewhat differently and at different rates. Therefore, we let $A(t)$ and $N(t)$ denote the population density of the

adult pea aphids and neonates, respectively, at time t. From previous studies, entomologists learned that the adult pea aphids reproduce best around 65 degrees Fahrenheit (18 degrees Celsius) and 80 percent of humidity [49]. Furthermore, each adult female aphid produces 6 to 8 neonates (nymphs) a day up to 100 nymphs. These reproductive characteristics of adult aphids can be incorporated into the birth rate $b(t, e, s)$ as a function of time t, environment e (temperature and humidity) and offspring productivity s. However, for simplicity, we here just assume a constant birthrate $b(t, e, s) = b$. In addition, one finds that it takes a newborn pea aphid or neonate τ days to become an adult aphid where τ can range from 5 to 7 days. Hence, a simple model for the pea aphids population density without the insecticides can be described by a distributed delay differential system

$$
\begin{aligned}
\dot{A}(t) &= \int_5^7 m(\tau)N(t - \tau)d\tau - d_A(t)A(t) \\
\dot{N}(t) &= bA(t) - \int_5^7 m(\tau)N(t - \tau)d\tau - d_N(t)N(t) \qquad (19) \\
N(s) &= 0, \qquad s \in [-7, 0), \\
N(0) &= N_0, \quad A(0) = 0,
\end{aligned}
$$

where $d_A(t)$ and $d_N(t)$ represent the natural death rate at time t of the adults and the neonates, respectively. Here $m(\tau)$ denotes the rate (probability) at which nymphs N become adults at age τ and move into the A compartment. Next we introduce the effects of the insecticide *Margosan-O* on the pea aphids into equation (19). It is known that *Margosan-O* can act on neonates and adults in several ways. First, the surfactant that causes the insecticide to adhere to plants' leaves can coat an aphid's exoskeleton upon contact, causing it to effectively smother to death shortly after exposure. Define $p_A(t)$ and $p_N(t)$ to be the probability of the adult and neonate aphids dying at time t due to *Margosan-O*. In the 1994 experiments [55], Stark and Rangus concluded that *Margosan-O* also affects the reproduction rate of the adult pea aphids. Thus, the birthrate b is no longer a constant but a function of insecticide concentration $\gamma \in [0, 100]$ for *Margosan-O* concentration ranges from 0 to 100 parts per million (ppm). The functional form of $b(\gamma)$ can be any decreasing function where $b(0) = b_{\max}$ and $b(100) = 0$. A further cause of death in the neonates is mutation and subsequent death during molting in the presence of insecticide. One manner in which one can represent this is to make the probability density p_N for death of neonates also dependent on the insecticide concentration γ. Of course, the adult death rate p_A also depends on γ. The aphids population density taking into account insecticide effects can then be modelled by the system of differential equations

$$
\begin{aligned}
\dot{A}(t) &= \int_5^7 m(\tau)N(t - \tau)d\tau - d_A(t)A(t) - p_A(t, \gamma)A(t) \\
\dot{N}(t) &= b(\gamma)A(t) - \int_5^7 m(\tau)N(t - \tau)d\tau - d_N(t)N(t) - p_N(t, \gamma)N(t) \qquad (20) \\
N(s) &= 0, \qquad s \in [-7, 0), \\
N(0) &= N_0, \quad A(0) = 0.
\end{aligned}
$$

This is a simple first model in which the quantities of interest include the probability densities p_A and p_N (which carry information about the sublethal effects of the insecticide) as well as the probability density m. It is easy to extend this model to more sophisticated and more useful models. For example, one might assume the birthrate is a probability density at time t over a continuum of pesticide concentrations γ. That is, $b(t, \gamma)$ is a probability density with corresponding time dependent distribution $dB(t, \gamma) = b(t, \gamma)d\gamma$ describing frequencies at which adult aphids give birth to neonates.

Our group is currently using models such as (20) with the data investigated in [3]; initial results are quite promising.

3   Theoretical Framework

In order to achieve our goal in a rigorous manner, we establish first existence and uniqueness results for equations (1) and (4), continuous dependence of the solution x on the parameter μ, and finally the fact that the directional derivative $y(t) = \delta x(t, \mu; \nu - \mu)$ exists and satisfies (4). Throughout this section, we will make use of the following form of the Gronwall's inequality, whose proof we do not present.

Lemma 1 *(Gronwall's inequality) Let $\varphi, \psi : [t_0, T] \to \mathbb{R}$ be two continuous and nonnegative functions and let K and L be positive constants, such that the inequality*

$$\varphi(t) \leq K + L \int_{t_0}^{t} \psi(s)\varphi(s)\,ds$$

holds for all $t \in [t_0, T]$. Then

$$\varphi(t) \leq K \exp\left(L \int_{t_0}^{t} \psi(s)\,ds \right)$$

for all $t \in [t_0, T]$.

Consider the abstract differential equation (1) where $f : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ is a continuous mapping in all three arguments; it is clear that for $t \geq t_0$, a solution $x(t, \mu)$ of (1) satisfies the equivalent integral equation

$$x(t, \mu) = x_0 + \int_{t_0}^{t} f(s, x(s, \mu), \mu)\,ds \qquad t \geq t_0. \tag{21}$$

Before studying the sensitivity of the solution of (1) with respect to the parameter μ, one must first establish that the solution of (1) exists and is unique. We define the *successive approximations* for the system (1) to be the functions $x^0, x^1, ...$, given recursively by

$$
\begin{aligned}
x^0(t, \mu) &= x_0, \\
x^{k+1}(t, \mu) &= x_0 + \int_{t_0}^{t} f(s, x^k(s, \mu), \mu)\,ds \qquad t \geq t_0,
\end{aligned}
$$

for $k = 0, 1, 2, ...$

Lemma 2 *(Existence and Uniqueness) Let $f : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ be continuous in t, x and μ and Lipschitz continuous in x, i.e.*

$$|f(t, x_1, \mu) - f(t, x_2, \mu)| \leq C|x_1 - x_2| \tag{22}$$

for some constant $C > 0$. Then the successive approximations x^k converge uniformly on $[t_0, T]$ to the unique solution x of (1).

Lemma 3 *(Continuous Dependence on Parameters) Let $f : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ be continuous in t, x and μ and Lipschitz continuous in x and for $\mu = \mu_0$ let $x(t, \mu_0)$ be a solution of (1) existing on $[t_0, T]$. Assume further that*

$$\lim_{\mu \to \mu_0} f(t, x, \mu) = f(t, x, \mu_0), \tag{23}$$

uniformly in (t, x). Then the differential equation (1) has a unique solution $x(t, \mu)$ satisfying

$$\lim_{\mu \to \mu_0} x(t, \mu) = x(t, \mu_0), \qquad t \in [t_0, T]. \tag{24}$$

The proofs of Lemma 2 and Lemma 3 above are omitted since they are very similar to those provided in [19] and in standard abstract differential equation results in the literature.

Theorem 1 *Let $f : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ of (1) be a continuous function in t, x and μ and have a Frechet derivative $f_x(t, x, \mu)$ with respect to x such that $f_x \in C[\mathbb{R}_+ \times X \times \mathcal{M}, B(X, X)]$, $|f_x(t, x, \mu)| \leq M_0$ for some constant $M_0 > 0$. Assume further that for $\mu, \nu \in \mathcal{M}$, f has a continuous directional derivative $\delta f(t, x, \mu; \nu - \mu)$ with respect to $\mu \in \mathcal{M}$ in the direction of $(\nu - \mu)$ such that $|\delta f(t, x, \mu; \nu - \mu)| \leq M_1$ where $M_1 > 0$. Then the directional derivative $y(t) = \delta x(t, \mu; \nu - \mu)$ with respect to μ exists, with $y : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ and satisfies the equation*

$$\begin{aligned} \dot{y}(t) &= f_x(t, x(t, \mu), \mu)y(t) + \delta f(t, x(t, \mu), \mu; \nu - \mu) \qquad \text{if } t \geq t_0, \\ y(t_0) &= 0. \end{aligned} \tag{25}$$

Proof: Since $f_x \in C[\mathbb{R}_+ \times X \times \mathcal{M}, B(X, X)]$ and $\delta f \in C[\mathbb{R}_+ \times X \times \mathcal{M}, X]$, by applying Lemma 2 we find that the differential equation (25) has a unique solution which we denote by $y(t)$. Let us consider the difference

$$m(t, \mu, \nu, \epsilon) = x(t, \mu + \epsilon(\nu - \mu)) - x(t, \mu)$$

for any two arbitrarily fixed $\mu, \nu \in \mathcal{M}$. Equivalently, we have

$$m(t, \mu, \nu, \epsilon) = \int_{t_0}^{t} \{f(s, x(s, \mu + \epsilon(\nu - \mu)), \mu + \epsilon(\nu - \mu)) - f(s, x(s, \mu), \mu)\}ds.$$

In order to show that the directional derivative of the solution x of (1) with respect to the parameter μ satisfies the sensitivity equation (25), we will prove that

$$\lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \left| \frac{x(t, \mu + \epsilon(\nu - \mu)) - x(t, \mu)}{\epsilon} - y(t) \right| = \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \left| \frac{m(t, \mu, \nu, \epsilon)}{\epsilon} - y(t) \right| = 0.$$

From the *Frechet* differentiability of f with respect to $x \in X$, we have

$$
\begin{aligned}
f(t, x(t, \mu+\epsilon(\nu - \mu)), \mu + \epsilon(\nu - \mu)) - f(t, x(t, \mu), \mu) &= f(t, x(t, \mu + \epsilon(\nu - \mu)), \mu + \epsilon(\nu - \mu)) \\
&\quad - f(t, x(t, \mu), \mu + \epsilon(\nu - \mu)) + f(t, x(t, \mu), \mu + \epsilon(\nu - \mu)) - f(t, x(t, \mu), \mu) \\
&= f_x(t, x(t, \mu), \mu + \epsilon(\nu - \mu))[x(t, \mu + \epsilon(\nu - \mu)) - x(t, \mu)] + w_1(m(t, \mu, \nu, \epsilon)) \\
&\quad + f(t, x(t, \mu), \mu + \epsilon(\nu - \mu)) - f(t, x(t, \mu), \mu)
\end{aligned}
$$

where

$$\frac{|w_1(m(t, \mu, \nu, \epsilon))|}{|m(t, \mu, \nu, \epsilon)|} \to 0$$

as $|m(t, \mu, \nu, \epsilon)| \to 0$. Consequently, we define $g_1(t, \mu, \nu, \epsilon)$ by

$$g_1(t, \mu, \nu, \epsilon) = \frac{|w_1(m(t, \mu, \nu, \epsilon))|}{|m(t, \mu, \nu, \epsilon)|}, \tag{26}$$

so that $g_1(t, \mu, \nu, \epsilon) \to 0$ uniformly in t as $\epsilon \to 0$. Now for $y(t)$ satisfying (25), we have

$$
\begin{aligned}
\left| \frac{m(t, \mu, \epsilon)}{\epsilon} - y(t) \right| &= \left| \int_{t_0}^{t} \left\{ f_x(s, x(s, t_0, x_0, \mu), \mu + \epsilon(\nu - \mu)) \frac{m(s, \mu, \epsilon)}{\epsilon} + \frac{w_1(m(s, \mu, \epsilon))}{\epsilon} \right. \right. \\
&\quad + \frac{f(s, x(s, t_0, x_0, \mu), \mu + \epsilon(\nu - \mu)) - f(s, x(s, t_0, x_0, \mu), \mu)}{\epsilon} \\
&\quad \left. \left. - f_x(s, x(s, t_0, x_0, \mu), \mu) y(s) - \delta f(s, x(s, t_0, x_0, \mu), \mu; \nu - \mu) \right\} ds \right|
\end{aligned}
$$

$$
\begin{aligned}
&\leq \int_{t_0}^{t} |f_x(s, x(s, \mu), \mu + \epsilon(\nu - \mu))| \left| \frac{m(s, \mu, \nu, \epsilon)}{\epsilon} - y(s) \right| ds + \int_{t_0}^{t} \frac{|w_1(m(s, \mu, \nu, \epsilon))|}{\epsilon} ds \\
&\quad + \int_{t_0}^{t} |f_x(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f_x(s, x(s, \mu), \mu)||y(s)| ds \\
&\quad + \int_{t_0}^{t} \left| \frac{f(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f(s, x(s, \mu), \mu)}{\epsilon} - \delta f(s, x(s, \mu), \mu; \nu - \mu) \right| ds.
\end{aligned}
$$

Since the Frechet derivative f_x is bounded and continuous with respect to the parameter μ, we have that

$$|f_x(s, x(s, \mu), \mu + \epsilon(\nu - \mu))| \leq M_0 \quad \text{for all } s \in [t_0, t],$$

14

and
$$g_2(s, \mu, \nu, \epsilon) = |f_x(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f_x(s, x(s, \mu), \mu)|$$

converges to 0 uniformly in s as $\epsilon \to 0$. In addition, we note that the solution $y(t)$ of equation (25) is continuous and therefore bounded by a constant $K_0 > 0$ for all $t \in [t_0, T]$. Also, from the definition of the directional derivative $\delta f(s, x(s, \mu), \mu; \nu - \mu)$, we have that

$$g_3(s, \mu, \nu, \epsilon) = \left| \frac{f(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f(s, x(s, \mu), \mu)}{\epsilon} - \delta f(s, x(s, \mu), \mu; \nu - \mu) \right|$$

converges in s to 0 as $\epsilon \to 0$. Using all the arguments above, we obtain that

$$\left| \frac{m(t, \mu, \nu, \epsilon)}{\epsilon} - y(t) \right| \leq \int_{t_0}^t M_0 \left| \frac{m(s, \mu, \nu, \epsilon)}{\epsilon} - y(s) \right| ds + \int_{t_0}^t \frac{|w_1(m(s, \mu, \nu, \epsilon))|}{\epsilon} ds$$

$$+ \int_{t_0}^t K_0 g_2(s, \mu, \nu, \epsilon) \, ds + \int_{t_0}^t g_3(s, \mu, \nu, \epsilon) \, ds.$$

Next we want to show that

$$\frac{|w_1(m(t, \mu, \nu, \epsilon))|}{\epsilon} \leq K_1 \frac{|w_1(m(t, \mu, \nu, \epsilon))|}{|m(t, \mu, \nu, \epsilon)|}$$

for some constant $K_1 > 0$. Hence, we want to consider

$$|m(t, \mu, \nu, \epsilon)| = \left| \int_{t_0}^t \left\{ f_x(s, x(s, \mu), \mu + \epsilon(\nu - \mu))[m(s, \mu, \nu, \epsilon)] + w_1(m(s, \mu, \nu, \epsilon)) \right. \right.$$

$$\left. \left. + f(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f(s, x(s, \mu), \mu) \right\} ds \right|$$

$$\leq \int_{t_0}^t \left\{ |f_x(s, x(s, \mu), \mu + \epsilon(\nu - \mu))||m(s, \mu, \nu, \epsilon)| + |w_1(m(s, \mu, \nu, \epsilon))| \right\} ds$$

$$+ \int_{t_0}^t |f(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f(s, x(s, \mu), \mu)| \, ds.$$

From equation(26), we obtain

$$|w_1(m(t, \mu, \nu, \epsilon))| = g_1(t, \mu, \nu, \epsilon)|m(t, \mu, \nu, \epsilon)|.$$

Using the definition of the directional derivative of $\delta f(t, x, \mu; \nu - \mu)$, we have that

$$|f(s, x(s, \mu), \mu + \epsilon(\nu - \mu)) - f(s, x(s, \mu), \mu)| \leq \epsilon g_3(s, \mu, \nu, \epsilon) + \epsilon |\delta f(s, x, \mu; \nu - \mu)|.$$

Furthermore, with the assumptions that $|f_x| \leq M_0$ and $|\delta f(t, x, \mu; \nu - \mu)| \leq M_1$, we find

$$|m(t, \mu, \nu, \epsilon)| \leq \int_{t_0}^t \left\{ [M_0 + g_1(s, \mu, \nu, \epsilon)]|m(s, \mu, \nu, \epsilon)| + \epsilon g_3(s, \mu, \nu, \epsilon) + |\epsilon \delta f(s, x, \mu; \nu - \mu)| \right\} ds$$

$$\leq \int_{t_0}^T \epsilon[g_3(s, \mu, \nu, \epsilon) + M_1] ds + \int_{t_0}^t [M_0 + g_1(s, \mu, \nu, \epsilon)]|m(s, \mu, \nu, \epsilon)| ds.$$

15

Applying Gronwall's inequality, we obtain that

$$|m(t, \mu, \nu, \epsilon)| \leq K_1 \epsilon,$$

where $K_1 = \left(\int_{t_0}^{T} (M_1 + g_3) ds \right) e^{\int_{t_0}^{T} (M_0 + g_1) ds}$ and $g_1(s, \mu, \nu, \epsilon)$ and $g_3(s, \mu, \nu, \epsilon)$ converge to 0 in s as $\epsilon \to 0$. It follows

$$\frac{|w_1(m(t, \mu, \nu, \epsilon))|}{\epsilon} \leq K_1 \frac{|w_1(m(t, \mu, \nu, \epsilon))|}{|m(t, \mu, \nu, \epsilon)|}.$$

Then we obtain

$$\left| \frac{m(t, \mu, \nu, \epsilon)}{\epsilon} - y(t) \right| \leq \int_{t_0}^{t} M_0 \left| \frac{m(s, \mu, \nu, \epsilon)}{\epsilon} - y(s) \right| ds + \int_{t_0}^{t} K_1 g_1(s, \mu, \nu, \epsilon) \, ds$$
$$+ \int_{t_0}^{t} K_0 g_2(s, \mu, \nu, \epsilon) \, ds + \int_{t_0}^{t} g_3(s, \mu, \nu, \epsilon) \, ds.$$

Applying Gronwall's inequality one more time, we have

$$\left| \frac{m(t, \mu, \nu, \epsilon)}{\epsilon} - y(t) \right| \leq \int_{t_0}^{T} [K_1 g_1(s, \mu, \nu, \epsilon) + K_0 g_2(s, \mu, \nu, \epsilon) + g_3(s, \mu, \nu, \epsilon)] \, ds \, e^{\int_{t_0}^{T} M_0 \, ds}$$

and taking the limit of $\epsilon \to 0$, we have

$$\lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} \left| \frac{m(t, \mu, \nu, \epsilon)}{\epsilon} - y(t) \right| = 0,$$

which completes the proof.

As noted in [19], under somewhat weaker assumptions of local Lipschitz conditions on f plus domination of f by an affine function one can establish similar results to those presented here.

4   Sensitivity Analysis for Probability Distribution Dependent Systems

We next apply the theoretical results established in the previous section to provide sensitivity equations for specific probability distribution dependent systems. We illustrate the ideas above with two examples possessing similarities to systems arising in modeling HIV cell dynamics [7] where a four-dimensional nonlinear distributed delay differential equation system is employed. Previous studies have focused on the theories and computations of sensitivity of the HIV model with respect to the distribution for separate cases of discrete distributions [6] and absolutely continuous (continuous) probability distributions or measures [19]. We

are interested in a theoretical and computational framework for the sensitivity with respect to a class of mixed probability distributions which have an absolutely continuous component and a saltus component in nonlinear distributed delay systems. Since we will be dealing with measures on real intervals, we recall that there is a one-to-one identification between the probability measures and their representations by probability distributions. Therefore, in general, we will not distinguish between a measure and its corresponding distribution when no confusion can occur. We first consider a scalar linear distributed delay differential equation and then a two-dimensional nonlinear distributed delay system. We present the theoretical framework for the sensitivity analysis in this section and discuss corresponding computations in the next section.

We consider the delay differential equation

$$
\begin{aligned}
\dot{x}(t) &= f(t) + \int_{-r}^{0} x(t+\tau) dP(\tau), \qquad t \geq 0 \\
x(\theta) &= \Phi(\theta), \quad \theta \in [-r, 0) \\
x(0) &= x^0,
\end{aligned}
\tag{27}
$$

where $x^0 \in \mathbb{R}$, $\Phi \in C(-r, 0)$ and $f \in L^2(0, T)$ for a fixed $T > 0$. We wish to analyze the sensitivity of the solution x with respect to the probability distribution P, which in this case is assumed to be the sum of an absolutely continuous part and a singular part with respect to the Lebesgue measure on the interval $S = [-r, 0]$, i.e.

$$
P(\tau) = \int_{-r}^{\tau} p(s)\, ds + \sum_{i=1}^{I} p_{\alpha_i} \Delta_{\alpha_i}(\tau), \quad \tau \in [-r, 0].
\tag{28}
$$

Here p is the density of the absolutely continuous part of P (guaranteed to exist a.e. by the Radon-Nikodym theorem) and Δ_{α_i} are Dirac measures with atoms at α_i, $i = 1, \ldots, I$ in the interval $[-r, 0]$. The main inconvenience we encounter when we try to investigate the sensitivity of x with respect to P is due to the fact that the set of the probability distributions of the form (28) is no longer a linear subspace of a Banach space and therefore we have to make sense of a meaningful way in which to perturb the probability distribution P and how to quantify the corresponding change in the solution x. As we shall see below, our general theoretical framework derived in Section 3 provides an elegant answer to this question if we interpret P as a "parameter" in an appropriately chosen space.

Let us denote by $(\mathcal{P}(S), \rho)$ the general space of probability measures defined on $S = [-r, 0]$ endowed with the Prohorov metric ρ, which can be formally defined as follows. For any closed subset $F \subset S$ and $\epsilon > 0$, we consider $F^\epsilon = \{q \in S : d(\tilde{q}, q) < \epsilon, \tilde{q} \in F\}$ and define the Prohorov metric $\rho : \mathcal{P}(S) \times \mathcal{P}(S) \to \mathbb{R}^+$ as

$$
\rho(P_1, P_2) \equiv \inf\{\epsilon > 0 : P_1[F] \leq P_2[F^\epsilon] + \epsilon, \ F \text{ closed}, \ F \subset S\}.
$$

Although the definition of ρ is not intuitive, one can show that

  (a) $(\mathcal{P}(S), \rho)$ is a complete metric space, and

(b) $\rho(P_k, P) \to 0$ as $k \to \infty$ is equivalent to

$$\int_S f\, dP_k(s) \to \int_S f\, dP(s)$$

as $k \to \infty$, for all bounded uniformly continuous $f : S \to \mathbb{R}$.

If we view $\mathcal{P}(S)$ as a subset of $\mathcal{X} = C_B^*(S)$, from the equivalence (b) above we obtain that convergence in the Prohorov metric is equivalent to the weak* convergence in $\mathcal{P}(S)$ [4, 23]. Here $C_B^*(S)$ denotes the topological dual of $C_B(S)$, the usual space of bounded continuous functions on S taken with the supremum norm and we take $\mathcal{X} = C_B^*(S)$ with the weak* topology. It is important to note that $\mathcal{X} = C_B^*(S)$ endowed with the weak* topology is a topological vector space which is not metrizable. However, the restriction of this topology to $\mathcal{P}(S)$ is metrizable and the metric which induces it is precisely the Prohorov metric ρ. In subsequent discussions in this paper, we shall take $\mathcal{P}(S)$ as the equivalent set of probability distributions on the finite interval $S = [-r, 0]$.

Next, let us denote by $\mathcal{M}$ the subset of $\mathcal{P}(S)$ of probability distributions P of the form (28), that is,

$$\mathcal{M} = \left\{ P \in \mathcal{P}(S) \mid P(\tau) = \int_{-r}^{\tau} p(s)\, ds + \sum_{i=1}^{I} p_{\alpha_i} \Delta_{\alpha_i}(\tau),\ I \in \mathbb{N} \right\}, \tag{29}$$

where $\mathbb{N}$ is the set natural numbers. Thus $\mathcal{M}$ is the set of probability distributions which can be expressed as the sum of an absolutely continuous part with respect to the Lebesgue measure and a singular part consisting of a finite sum of Dirac measures. Using the definition (29), one can easily check that $\mathcal{M}$ is a convex subset of $(\mathcal{P}(S), \rho)$ and therefore a convex subset of $C_B^*(S)$. If we take the convex set $\mathcal{M}$ defined above with the Prohorov metric, we can apply the previous theory developed in Section 3 to perform the sensitivity analysis of (27) with respect to P.

Let $(x(t), x_t) \in X = \mathbb{R} \times C(-r, 0; \mathbb{R})$ and $\mathcal{M}$ be given by (29); then (27) can be reformulated as an abstract system of the form (5) where the right side $F : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ is defined as

$$F(t, x, x_t, P) = f(t) + \int_{-r}^{0} x_t(\tau)\, dP(\tau), \tag{30}$$

which is precisely that of the linear delay equation (27). In order to quantify how the solution of the equation (27) changes with respect to variations in the probability distribution P, we fix another probability distribution Q in the space $\mathcal{M}$ and show that the *sensitivity* $y = \frac{\partial x}{\partial P}(t, P; Q - P)$ *of x with respect to P* defined as the directional derivative of x at P in the direction $Q - P$ exists and is uniquely defined. Then, by the theory of the previous section, $y : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ satisfies

$$\begin{aligned} \dot{y}(t) &= \frac{\partial F}{\partial x}(t, x, x_t, P) y(t) + \frac{\partial F}{\partial x_t}(t, x, x_t, P) y_t + \frac{\partial F}{\partial P}(t, x, x_t, P; Q - P), \\ y(\tau) &= 0, \quad \tau \in [-r, 0], \end{aligned} \tag{31}$$

where F is defined in (30). In the equation above x_t represents the *delay function* defined on $[-r, 0]$ and given by $x_t(\tau) = x(t + \tau)$. Similarly, y_t represents the *delay sensitivity function* given by $y_t(\tau) = y(t + \tau)$ for $-r \leq \tau \leq 0$. The quantities $\partial F/\partial x$, $\partial F/\partial x_t$ and $\partial F/\partial P$ represent the *Frechet derivatives* of F with respect to x, x_t and, respectively, the directional derivative of F at P in the direction $Q - P$. It is easy to verify that $\partial F/\partial x = 0$, and because the integral appearing in (30) is linear in x_t, the Frechet derivative of F with respect to x_t is simply

$$\frac{\partial F}{\partial x_t}(t, x, x_t, P)[y_t] = \int_{-r}^{0} y_t(\tau) dP(\tau). \tag{32}$$

In order to evaluate the directional derivative $\partial F/\partial P$, let $Q \in \mathcal{M}$ be another probability distribution of the form

$$Q(\tau) = \int_{-r}^{\tau} q(s)\, ds + \sum_{j=1}^{J} q_{\beta_j} \Delta_{\beta_j}(\tau).$$

Then the partial derivative $\partial F/\partial P$ at P in the direction $Q - P$ is well defined and given by

$$\frac{\partial F}{\partial P}(t, x, x_t, P; Q - P) = \lim_{\varepsilon \to 0} \frac{F(t, x, x_t, P + \varepsilon(Q - P)) - F(t, x, x_t, P)}{\varepsilon}$$
$$= \lim_{\varepsilon \to 0} \frac{\int_{-r}^{0} x(t + \tau) d[P + \varepsilon(Q - P)](\tau) - \int_{-r}^{0} x(t + \tau) dP(\tau)}{\varepsilon}.$$

Given the form of $P + \varepsilon(Q - P)$ and P, we obtain that

$$\frac{\partial F}{\partial P}(t, x, x_t, P; Q - P) = \lim_{\varepsilon \to 0} \frac{\int_{-r}^{0} x(t + \tau)[p(\tau) + \varepsilon(q(\tau) - p(\tau))]d\tau - \int_{-r}^{0} x(t + \tau)p(\tau)d\tau}{\varepsilon}$$
$$+ \lim_{\varepsilon \to 0} \frac{\sum_{i=1}^{I}(1 - \varepsilon)p_{\alpha_i}x_t(\tau_{\alpha_i}) + \sum_{j=1}^{J} \varepsilon q_{\beta_j}x_t(\tau_{\beta_j}) - \sum_{i=1}^{I} p_{\alpha_i}x_t(\tau_{\alpha_i})}{\varepsilon}$$
$$= \int_{-r}^{0} x_t(\tau)[q(\tau) - p(\tau)]d\tau + \sum_{j=1}^{J} q_{\beta_j}x_t(\tau_{\beta_j}) - \sum_{i=1}^{I} p_{\alpha_i}x_t(\tau_{\alpha_i}). \tag{33}$$

With F given in (30) and $\partial F/\partial x_t$ and $\partial F/\partial P$, respectively, derived in equation (32) and equation (33), Theorem 1 provides the existence and uniqueness of $y(t)$ defined in (31) above. More explicitly, the sensitivity $y(t) = \frac{\partial x}{\partial P}(t, P; Q - P)$ is well defined and satisfies the delay differential equation

$$\dot{y}(t) = \int_{-r}^{0} y_t(\tau) dP(\tau) + \int_{-r}^{0} x_t(\tau)[q(\tau) - p(\tau)]d\tau + \sum_{j=1}^{J} q_{\beta_j}x_t(\tau_{\beta_j}) - \sum_{i=1}^{I} p_{\alpha_i}x_t(\tau_{\alpha_i}), \tag{34}$$

$$y(\tau) = 0, \quad \tau \in [-r, 0].$$

We next consider a nonlinear example

$$
\begin{aligned}
\dot{x}_1(t) &= f(t) + x_2(t)\int_{-r}^{0} x_1(t+\tau)dP(\tau) \\
\dot{x}_2(t) &= g(t) + ax_1(t) + bx_2(t) \\
x_1(\theta) &= \Phi(\theta), \quad \theta \in [-r, 0) \\
x_1(0) &= c_1 \\
x_2(0) &= c_2,
\end{aligned}
\tag{35}
$$

where $c_1, c_2 \in \mathbb{R}, \Phi \in C(-r,0)$ and $f, g \in L^2(0,T)$ for fixed $T > 0$. We define $X = \mathbb{R}^2 \times C(-r,0;\mathbb{R}^2)$ and let $(x(t), x_t) \in X$ where $x(t) = (x_1(t), x_2(t))^T$. Then for $\mathcal{M}$ defined in (29), the system (35) can be written in the vector form

$$
\begin{aligned}
\dot{x}(t) &= F(t, x, x_t, P) \tag{36} \\
(x(0), x_0) &= ((c_1, c_2)^T, (\Phi, 0)^T), \tag{37}
\end{aligned}
$$

where $x_t(\tau) = x(t+\tau)$ and $F : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ is a vector function whose components are given by the right sides of (35). Again, applying Theorem 1 from Section 3, we find that the sensitivity $y : \mathbb{R}_+ \times X \times \mathcal{M} \to X$ given by $y(t) = (\partial x/\partial P)(t, P; Q-P)$ of system (36) satisfies

$$
\begin{aligned}
\dot{y}(t) &= \frac{\partial F}{\partial x}(t, x, x_t, P)y(t) + \frac{\partial F}{\partial x_t}(t, x, x_t, P)[y_t] + \frac{\partial F}{\partial P}(t, x, x_t, P; Q-P), \\
y(\tau) &= 0, \quad \tau \in [-r, 0]. \tag{38}
\end{aligned}
$$

Using the definition of the Frechet derivative we can easily establish that

$$
\frac{\partial F}{\partial x}(t, x, x_t, P)y(t) = \begin{bmatrix} y_2(t)\int_{-r}^{0} x_1(t+\tau)dP(\tau) \\ ay_1(t) + by_2(t) \end{bmatrix},
$$

$$
\frac{\partial F}{\partial x_t}(t, x, x_t, P)[y_t] = \begin{bmatrix} x_2(t)\int_{-r}^{0} y_1(t+\tau)dP(\tau) \\ 0 \end{bmatrix},
$$

and using the definition of the directional derivative of F at P in the direction $Q-P$, we have

$$
\frac{\partial F}{\partial P}(t, x, x_t, P; Q-P) =
$$
$$
= \begin{bmatrix} x_2(t)\left(\int_{-r}^{0} x_1(t+\tau)[q(\tau)-p(\tau)]d\tau + \sum_{j=1}^{J} q_{\beta_j}x_1(t+\tau_{\beta_j}) - \sum_{i=1}^{I} p_{\alpha_i}x_1(t+\tau_{\alpha_i})\right) \\ 0 \end{bmatrix}.
$$

By substituting these derivatives in (38) we obtain the sensitivity equation for (35).

5 Examples

In this section we present numerical results for the two examples given in the previous section. We outline the scheme to approximate the state solution and the solution of the corresponding sensitivity equation and then illustrate the convergence of the scheme with numerical simulations.

5.1 A Linear Example

We first discuss numerical simulations for the linear delay differential equation defined in (27) and the corresponding sensitivity equation (34). The two equations can be written together as a coupled system

$$
\dot{x}(t) = f(t) + \sum_{i=1}^{I} p_{\alpha_i} x(t + \tau_{\alpha_i}) + \int_{-r}^{0} x(t + \tau) p(\tau) \, d\tau, \qquad t \geq 0
$$

$$
\dot{y}(t) = \sum_{i=1}^{I} p_{\alpha_i} y(t + \tau_{\alpha_i}) + \sum_{j=1}^{J} q_{\beta_j} x(t + \tau_{\beta_j}) - \sum_{i=1}^{I} p_{\alpha_i} x(t + \tau_{\alpha_i}) + \int_{-r}^{0} y(t + \tau) p(\tau) \, d\tau
$$

$$
+ \int_{-r}^{0} x(t + \tau)[q(\tau) - p(\tau)] d\tau \tag{39}
$$

$$
\begin{aligned}
x(\theta) &= \Phi(\theta), & y(\theta) = 0, & \qquad \theta \in [-r, 0), \\
x(0) &= x^0, & y(0) = 0,
\end{aligned}
$$

where $\tau_{\alpha_i}, i = 1, \ldots, I$, and $\tau_{\beta_j}, j = 1, \ldots, J$, are atoms in $[-r, 0]$ for the associated Dirac measures. In order to define approximations for (39), we first rewrite the system as an operator equation in a Hilbert space Z (e.g., see [10, 13]). We let $v = (x, y)^T$ and $z(t) = (v(t), v_t) \in Z \equiv \mathbb{R}^2 \times L^2(-r, 0; \mathbb{R}^2)$, where v_t is the function $\tau \rightarrow v(t + \tau)$ for $\tau \in [-r, 0)$. Given $\eta = (\psi^0, \zeta^0)^T \in \mathbb{R}^2$ and $\varphi = (\psi, \zeta)^T \in C(-r, 0; \mathbb{R}^2)$, we define a linear operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset Z \rightarrow Z$ with domain $\mathcal{D}(\mathcal{A}) = \{(\eta, \varphi) \in Z \mid \varphi \in H^1(-r, 0; \mathbb{R}^2) \text{ and } \eta = \varphi(0)\}$ by $\mathcal{A}(P, Q)[\eta, \varphi] = (L(P, Q)[\varphi], \dot{\varphi})$ where

$$
L(P, Q)[\varphi] = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \sum_{i=1}^{I} p_{\alpha_i} \varphi(t + \tau_{\alpha_i}) \quad + \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \sum_{j=1}^{J} q_{\beta_j} \varphi(t + \tau_{\beta_j})
$$

$$
+ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \int_{-r}^{0} \varphi(\tau) p(\tau) \, d\tau \quad + \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \int_{-r}^{0} \varphi(\tau)[q(\tau) - p(\tau)] \, d\tau.
$$

Then the delay system (39) above, when formulated as an abstract Cauchy problem, is given by

$$
\begin{aligned}
\dot{z}(t) &= \mathcal{A}(P, Q) z(t) + G(t) \\
z(0) &= z_0,
\end{aligned} \tag{40}
$$

where $G(t) = (f(t), 0)^T$ and $z_0 = ((x^0, 0)^T, (\Phi, 0)^T)$.

Under the assumption that $P, Q \in \mathcal{M}$ are known and fixed, the solution z of (40) can be approximated using either the averaging scheme developed by Banks and Burns in [10] or the spline scheme developed by Banks and Kappel in [13]. Here we chose the linear spline approximation scheme for our computations and hence we briefly discuss the resulting numerical framework. Following [13], we define Z^N to be an approximating piecewise linear spline subspace of Z, Π^N as the orthogonal projection of Z onto Z^N, and $\mathcal{A}^N(P, Q)$ as the approximating operator for $\mathcal{A}(P, Q)$ given by $\mathcal{A}^N(P, Q) = \Pi^N \mathcal{A}(P, Q) \Pi^N$. Then the Cauchy problem (40) is approximated by the finite dimensional Cauchy problem

$$
\begin{aligned}
\dot{z}^N(t) &= \mathcal{A}^N(P, Q) z^N(t) + \Pi^N G(t), \qquad t \geq 0 \\
z^N(0) &= \Pi^N z_0.
\end{aligned}
\tag{41}
$$

It follows from arguments in [13] that for fixed $(P, Q) \in \mathcal{M}^2$, the approximating scheme $\{Z^N, \Pi^N, \mathcal{A}^N(P, Q)\}$ yields solutions $z^N(t)$ which converge to the solution $z(t)$ of (40) uniformly in t on any finite interval as $N \to \infty$. For the linear spline approximation, we fix the basis for a special case Z_1^N of Z^N corresponding to the partition of $[-r, 0]$ by $t_k^N = -k(r/N)$ for $k = 0, 1, ..., N$. Then the bases are defined by $\hat{\beta}^N = (\beta^N(0), \beta^N)$ for $\beta^N = (e_0^N, e_1^N, ..., e_N^N) \bigotimes I_n$ where I_n denotes the $n \times n$ identity matrix and the piecewise linear e_l^N's are defined by

$$
e_l^N(t_k^N) = \delta_{lk} \qquad \text{for} \quad l, k = 0, 1, .., N.
$$

When $\mathcal{A}^N(P, Q)$ is restricted to Z_1^N, we have a matrix representation $A_1^N(P, Q)$ of $\mathcal{A}^N$. Next we define $w^N(t)$ and $G^N(t)$ so that $z^N(t) = \hat{\beta}^N w^N(t)$ and $\Pi^N G(t) = \hat{\beta}^N G^N(t)$, respectively. Then solving for $z^N(t)$ in system (41) is equivalent to solving for $w^N(t)$ in the linear ordinary differential equation (ODE)

$$
\begin{aligned}
\dot{w}^N(t) &= A_1^N(P, Q) w^N(t) + G^N(t), \qquad t \geq 0 \\
w^N(0) &= w_0^N,
\end{aligned}
\tag{42}
$$

where $\hat{\beta}^N w_0^N = \Pi^N z_0$. We refer the reader to [13] for more details on convergence theory for the spline approximations and precise details for construction of $A_1^N(P, Q)$, $\Pi^N z_0$, and $\Pi^N G$. In general, application of the linear spline scheme to an n dimensional delay system is equivalent to solving an $n(N + 1)$ dimensional linear ODE system. One can think of the solutions w^N as the generalized Fourier coefficients when one expands the solution z using $(N + 1)$ piecewise linear splines as basis elements.

We next define sets of approximating probability distributions by

$$
\mathcal{M}^M = \left\{ P^M \in \mathcal{P} \mid P^M(\tau) = \int_{-r}^{\tau} \sum_{k=0}^{M} a_k^M l_k^M(s) ds + \sum_{i=1}^{I} p_{\alpha_i} \Delta_{\alpha_i}(\tau) \right\},
$$

22

where the l_j^M's are the usual piecewise linear splines on $[-r, 0]$. It is known (see [14]) that for any P with corresponding absolutely continuous part p, one can, by appropriate choice of the coefficients $\{a_k^M\}$, obtain convergence of $p^M = \sum a_k^M l_k^M$ to p in L^2. This guarantees convergence of P^M to P in the Prohorov metric. We note that any approximation scheme that guarantees weak L^2 convergence of $p^M \to p$ could be used in the approximations P^M for P in the Prohorov sense (e.g., see [20]).

For a given P, Q suppose that P^M, $Q^M \in \mathcal{M}^M$ have been chosen so that $P^M \to P$, $Q^M \to Q$ in the Prohorov sense. Then we approximate the system (40) by

$$
\begin{aligned}
\dot{z}^{N,M}(t) &= \mathcal{A}^N(P^M, Q^M) z^{N,M}(t) + \Pi^N G(t), \qquad t \geq 0 \\
z^{N,M}(0) &= \Pi^N z_0.
\end{aligned}
\tag{43}
$$

The convergence of $z^{N,M}$ to the solution z of (40) as $N, M \to \infty$ can be argued using standard convergence arguments as in [14] and the convergence $z^N \to z$, $P^M \to P$ and $Q^M \to Q$ as $N, M \to \infty$. We illustrate this convergence with simulations for a specific case of the linear delay system (27) for special choices of the distributions P and Q.

Let

$$
dP(\tau) = \frac{1}{4} \delta_{-r}(\tau) d\tau + \frac{3}{4d} p(\tau) d\tau,
\tag{44}
$$

where $p(\tau)$ is a Gaussian probability density

$$
p(\tau) = \frac{1}{\sigma\sqrt{(2\pi)}} e^{-\frac{(\tau - \bar{\tau})^2}{2\sigma^2}}
\tag{45}
$$

with $\bar{\tau} = -0.35$, $\sigma^2 = 0.1$, $r = 1$ and $d = \int_{-r}^0 p(s) ds$. We take Q to be a simple uniform distribution on $[-1, 0]$ so that $dQ(\tau) = q(\tau) d\tau = 1 d\tau$, $\tau \in [-1, 0]$. We choose $f(t) = 1$, $\Phi(\theta) = 0$, and $x^0 = 1$. Then the coupled system (39) for x and the corresponding sensitivity y in the direction $Q - P$ for this specific situation is given by

$$
\begin{aligned}
\dot{x}(t) &= 1 + \frac{1}{4} x(t-1) + \frac{3}{4d} \int_{-1}^0 x(t+\tau) p(\tau) d\tau \\
\dot{y}(t) &= \frac{1}{4} y(t-1) - \frac{1}{4} x(t-1) + \frac{3}{4d} \int_{-1}^0 y(t+\tau) p(\tau) d\tau + \int_{-1}^0 x(t+\tau)[1 - \frac{3}{4d} p(\tau)] d\tau \\
x(\theta) &= y(\theta) = 0, \qquad \theta \in [-1, 0), \\
x(0) &= 1, \qquad y(0) = 0.
\end{aligned}
\tag{46}
$$

We first fix $N = 16$ and solve the approximation system (43) for different values of M and plot the approximations $x^{N,M}$, $y^{N,M}$ of the solution x and the corresponding sensitivity $y(t)$, respectively, in Figure 3. These graphs illustrate the convergence guaranteed by arguments similar to those in [14]; in Figure 3(a) we observe that $x^{N,M} \to x$ and in Figure 3(b) that

23

$y^{N,M} \to y$ for a fixed $N = 16$ and M increasingly large. We then fix $M = 16$ and solve system (43) for different N values. The approximating solutions $x^{N,M}$ and $y^{N,M}$ for a fixed $M = 16$ and N increasingly large are graphed in Figure 4. Again, one can see from Figure 4 that we have convergence for both the solutions $x^{N,M}$ and sensitivity $y^{N,M}$ as $N \to \infty$ and $M = 16$. We remark that by solving here the approximation to the system (46), we approximate x and y simultaneously. Alternatively, one can approximate x and y separately by first approximating x to obtain $x^{N,M}$ and then using $x^{N,M}$ in solving the equation for $y^{N,M}$.

(a) (b)
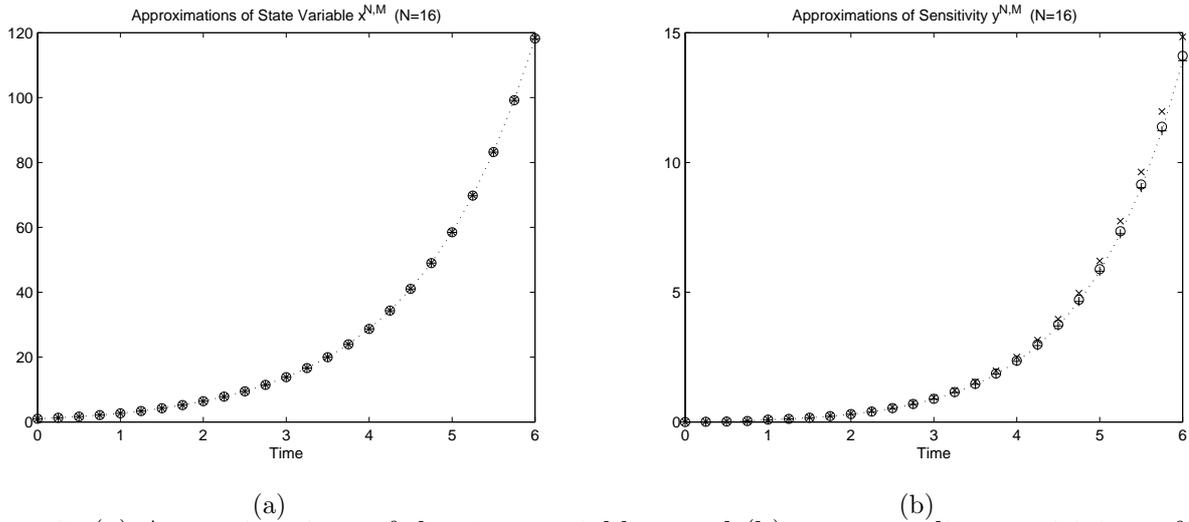
Figure 3: (a) Approximations of the state variable x and (b) corresponding sensitivity y for fixed $N = 16$ and varying M=4 (cross), M=8 (circle), M=16 (plus), M=32 (dotted line).

(a) (b)
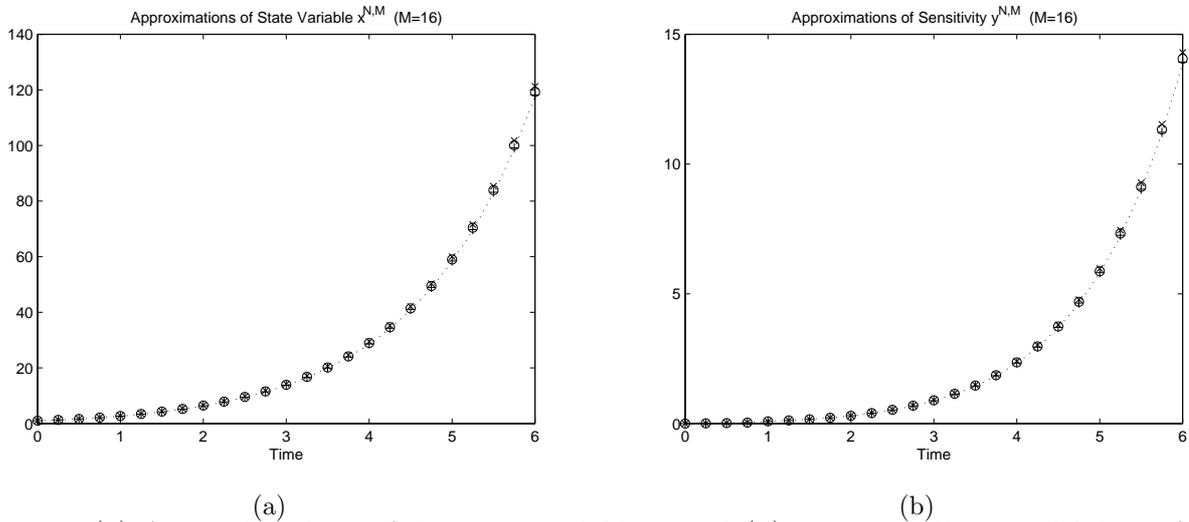
Figure 4: (a) Approximations of the state variable x and (b) corresponding sensitivity y for fixed $M = 16$ and varying N=4 (cross), N=8 (circle), N=16 (plus), N=32 (dotted line).

24

5.2   A Nonlinear Example

We consider next the nonlinear delay differential system (35) and the corresponding sensitivity system (38) derived in Section 4. Upon substituting for P, Q and coupling the delay system (35) and its sensitivity equation (38), we obtain

$$
\begin{aligned}
\dot{x}_1(t) &= f(t) + x_2(t) \sum_{i=1}^{I} p_{\alpha_i} x_1(t + \tau_{\alpha_i}) + x_2(t) \int_{-r}^{0} x_1(t + \tau) p(\tau) \, d\tau, \qquad t \geq 0 \\
\dot{x}_2(t) &= g(t) + a x_1(t) + b x_2(t) \\
\dot{y}_1(t) &= x_2(t) \left(\sum_{i=1}^{I} p_{\alpha_i} y_1(t + \tau_{\alpha_i}) + \sum_{j=1}^{J} q_{\beta_j} x_1(t + \tau_{\beta_j}) \right) + y_2(t) \int_{-r}^{0} x_1(t + \tau) p(\tau) d\tau \\
&\quad + x_2(t) \left[\int_{-r}^{0} y_1(t + \tau) p(\tau) \, d\tau + \int_{-r}^{0} x_1(t + \tau)[q(\tau) - p(\tau)] d\tau \right] \\
&\quad + (y_2(t) - x_2(t)) \left[\sum_{i=1}^{I} p_{\alpha_i} x_1(t + \tau_{\alpha_i}) \right] \\
\dot{y}_2(t) &= a y_1(t) + b y_2(t) \\
x_1(\theta) &= \Phi(\theta), \quad y_1(\theta) = 0, \qquad \theta \in [-r, 0) \\
x_i(0) &= x_i^0, \ \ y_i(0) = 0, \ \ i = 1, 2,
\end{aligned}
\tag{47}
$$

where τ_{α_i} and $\tau_{\beta_i} \in [-r, 0]$. We formulate an abstract Cauchy problem for system (47) by employing $\eta = (\psi_1^0, \psi_2^0, \zeta_1^0, \zeta_2^0)^T \in \mathbb{R}^4$, $\varphi = (\psi_1, \psi_2, \zeta_1, \zeta_2)^T \in C(-r, 0; \mathbb{R}^4)$, $v = (x_1, x_2, y_1, y_2)^T \in \mathbb{R}^4$, and $z = (v(t), v_t) \in Z = \mathbb{R}^4 \times L^2(-r, 0; \mathbb{R}^4)$. We may then define a nonlinear operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset Z \to Z$ and $\mathcal{D}(\mathcal{A}) = \{(\eta, \varphi) \in Z \,|\, \varphi \in H^1(-r, 0; \mathbb{R}^4) \text{ and } \eta = \varphi(0)\}$. For $(\eta, \varphi) \in \mathcal{D}(\mathcal{A})$, define $\mathcal{A}(P, Q)[\eta, \varphi] = (L(\eta) + F_1(P, Q)[\eta, \varphi], \dot{\varphi})$ where

$$
L(\eta) =
\begin{bmatrix}
0 & 0 & 0 & 0 \\
a & b & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & a & b
\end{bmatrix}
\eta
$$

and

$$
F_1(P, Q)[\eta, \varphi] = C_1 \sum_{i=1}^{I} p_{\alpha_i} \eta_2 \varphi(t + \tau_{\alpha_i}) \ + \ D_{(3,1)} \sum_{j=1}^{J} q_{\beta_j} \eta_2 \varphi(t + \tau_{\beta_j}) \ + \ C_1 \int_{-r}^{0} \eta_2 p(\tau) \varphi(\tau) \, d\tau
$$

$$
+ D_{(3,1)} \sum_{i=1}^{I} p_{\alpha_i} \eta_4 \varphi(t + \tau_{\alpha_i}) + D_{(3,1)} \left(\int_{-r}^{0} \eta_2 \varphi(\tau) q(\tau) \, d\tau + \int_{-r}^{0} \eta_4 \varphi(\tau) p(\tau) \, d\tau \right),
$$

where $D_{(i,j)}$ is a 4×4 matrix with 1 at the $(i^{\text{th}}, j^{\text{th}})$ entry and zero everywhere else and

$$
C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
$$

Then the abstract Cauchy formulation of the distributed delay system (47) above is

$$
\begin{aligned}
\dot{z}(t) &= \mathcal{A}(P,Q)z(t) + G(t) \\
z(0) &= z_0,
\end{aligned} \tag{48}
$$

where $G(t) = ((f(t), g(t), 0, 0)^T, (0, 0, 0, 0)^T)$ and $z_0 = ((x_1^0, x_2^0, 0, 0)^T, (\Phi, 0, 0, 0)^T)$. For numerical purposes one can employ an approximation framework similar to that in the linear example above; we do not repeat details here. We note that the combined results in ([2],[7],[13]) can be used to establish that the resulting piecewise linear spline approximating scheme $\{Z^N, \Pi^N, \mathcal{A}^N(P,Q)\}$ provides convergence of $z^N(t)$ to $z(t)$ for the system (47). Here $z^N(t)$ is the solution of the corresponding approximation system similar to system (41) given above for the linear example. We again may approximate the distributions P, Q in the Prohorov sense so that $P^M \to P$ and $Q^M \to Q$ as $M \to \infty$, and again construct corresponding approximating systems similar to system (43) in the linear example. In this case, one also can argue the convergence of the approximations $z^{N,M}(t) \to z(t)$ as $N, M \to \infty$ using $z^N(t) \to z(t)$ for $N \to \infty$ and $P^M \to P$, $Q^M \to Q$ as $M \to \infty$.

We again illustrate convergence with simulations corresponding to specific distributions P and Q. In particular, we again choose P to be given in (44) and Q to be the uniform distribution $dQ(\tau) = q(\tau)\, d\tau = 1\, d\tau$ for $\tau \in [-1, 0]$. We choose $f(t) = 1$, $g(t) = 2$, $a = -3$ and $b = -2$ and define $\Phi(\theta) = 0$ for $\theta \in [-1, 0)$ and $x_1^0 = x_2^0 = 1$. Then we solve the corresponding approximating system which is similar to the system given in (43); we first fix $N = 16$ and let M vary. The approximations of the solution $x^{N,M}(t)$ are plotted in Figure 5 and the sensitivity $y^{N,M}(t)$ are depicted in Figure 6. We see from Figure 5 and Figure 6 that with a fixed $N = 16$ and $M \to \infty$, both $x^{N,M}$ and the corresponding sensitivity $y^{N,M}$ converge. We then take M fixed at 16 and let N vary. Similar to the previous case where N is fixed, we observe from Figure 7 and Figure 8 where we plot $x^{N,M}$ and $y^{N,M}$ respectively, that $x^{N,M}$ and $y^{N,M}$ converge to x and y for $M = 16$ and $N \to \infty$.
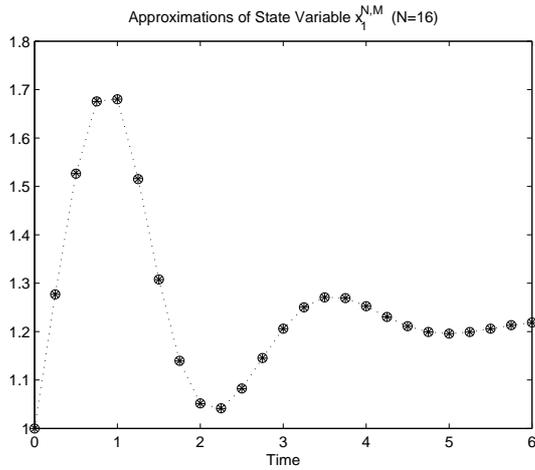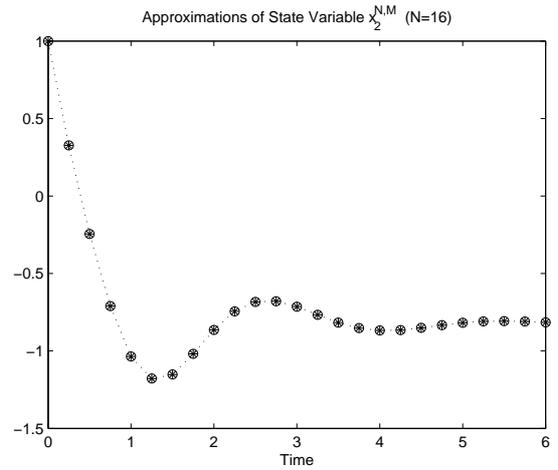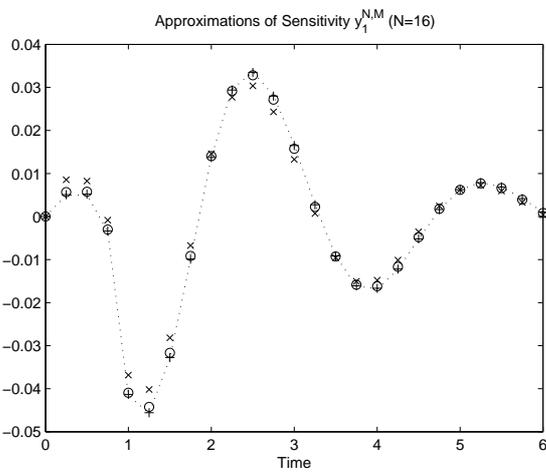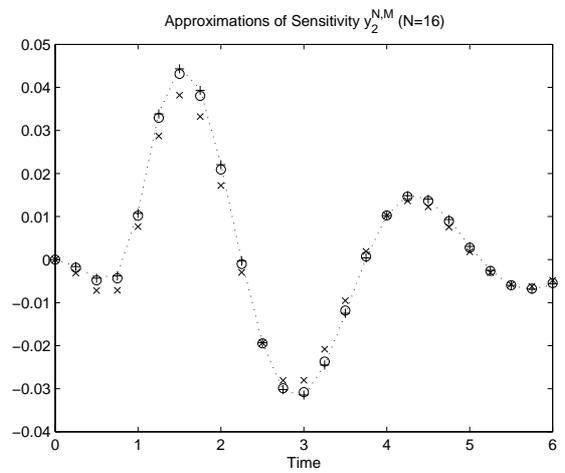
Figure 5: (a) Approximations of the state variable x_1 and (b) approximations of the state variable x_2 for fixed $N = 16$ and varying M=4 (cross), M=8 (circle), M=16 (plus) and M=32 (dotted line).

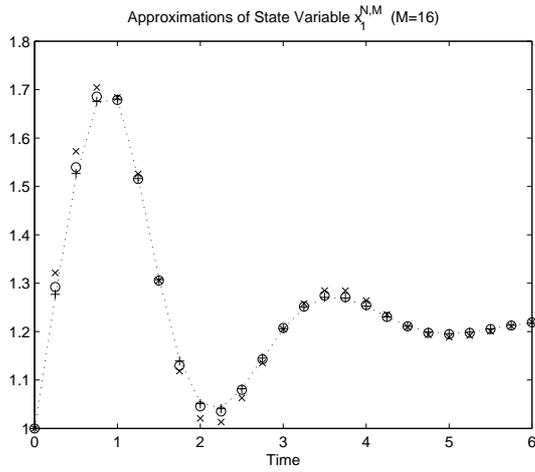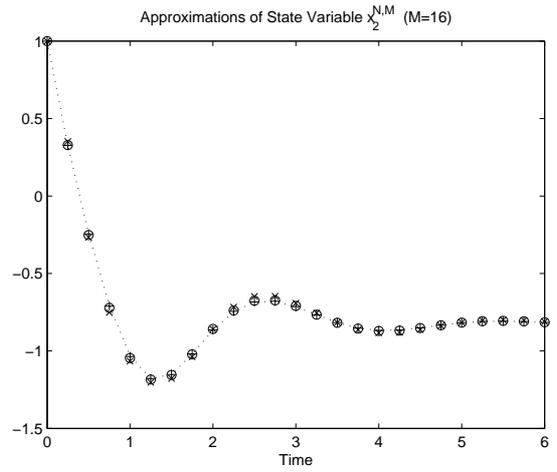Figure 6: (a) Approximations of y_1, the sensitivity of x_1 and (b) approximations of y_2, the sensitivity of x_2 for fixed $N = 16$ and varying M=4 (cross), M=8 (circle), M=16 (plus) and M=32 (dotted line).
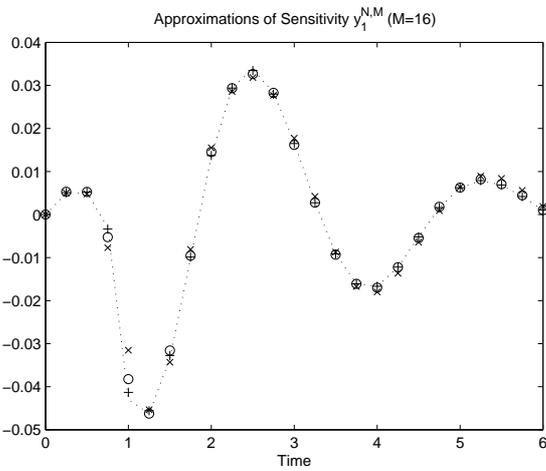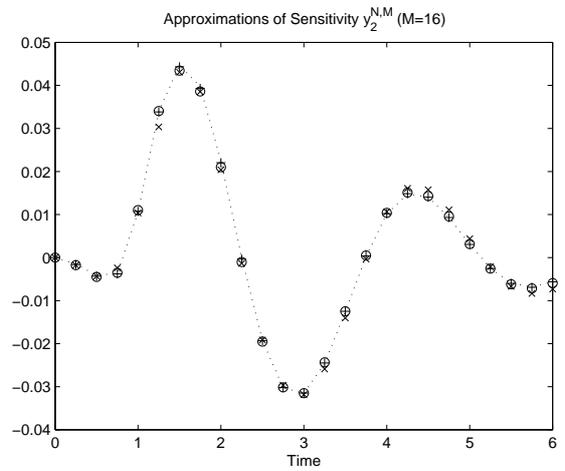
Figure 7: (a) Approximations of the state variable x_1 and (b) approximations of the state variable x_2 for fixed $M = 16$ and varying N=4 (cross), M=8 (circle), N=16 (plus), N=32 (dotted line).

Figure 8: (a) Approximations of y_1, the sensitivity of x_1 and (b) approximations of y_2, the sensitivity of x_2 for fixed $M = 16$ and varying N=4 (cross), M=8 (circle), N=16 (plus), N=32 (dotted line).

6   Concluding Remarks

In this paper we have developed both theoretical and computational aspects of a sensitivity methodology for probability measure dependent dynamical systems. In addition to the usual uses of sensitivities, the results developed here represent a first step toward a general mathematical and statistical methodology for estimation of uncertainties (in this case represented by probability distributions P) in the dynamics of a system. In particular, we refer to a theory for the asymptotic properties of ordinary least squares (OLS) estimators $\hat{P}_{OLS}$ which could lead to standard errors and confidence "intervals" in a functional setting. In the standard finite dimensional parameter theory (see [35] and Chapter 12 of [52]), the asymptotic theory is given in terms of the covariance matrix Σ^2 which can be approximated by

$$\Sigma^2 \approx \sigma_0^2 (\chi^T \chi)^{-1},$$

where χ is the sensitivity matrix

$$\chi = \left(\frac{\partial x_i}{\partial \theta_j} \right).$$

For the systems of interest here, the goal is to develop a rigorous mathematical/statistical framework for OLS estimators for "parameters" $P \in \mathcal{P}(S)$, i.e., the parameters to be estimated are actually probability distributions in an infinite dimensional set. For this an operator analogue of the covariance matrix will involve the sensitivities

$$\chi = \left(\frac{\partial x}{\partial P} \right).$$

Thus, in addition to use in routine sensitivity analysis for uncertainty in dynamical systems, the theory presented above provides a foundation to develop an operator theoretic form of the needed asymptotic distribution theory similar to the one currently available for estimation in a finite dimensional parameter setting.

References

[1] H. M. ADELMAN AND R.T. HAFTKA, Sensitivity analysis of discrete structural systems, *A.I.A.A. Journal*, **24** (1986), 823–832.

[2] H.T. BANKS, Approximation of nonlinear functional differential equation control systems, *J. Optimization Theory and Applications*, **29** (1979), 383–408.

[3] H.T. BANKS, J.E. BANKS, L.K. DICK AND J.D. STARK, Estimation of dynamic rate parameters in insect populations undergoing sublethal exposure to pesticides, *CRSC-TR05-22*, NCSU, May, 2005; *Bulletin of Mathematical Biology*, submitted.

[4] H.T. BANKS AND K. L. BIHARI, Modeling and estimating uncertainty in parameter estimation, *CRSC-TR99-40*, NCSU, August, 1999; *Inverse Problems*, **17** (2001), 1-17.

[5] H.T. BANKS, V.A. BOKIL, S. HU, A.K. DHAR, R.A. BULLIS, C.L. BROWDY AND F.C.T. ALLNUTT, Shrimp biomass and viral infection for production of biological countermeasures. *CRSC-TR05-45*, NCSU, December, 2005; *Math. Biosci. and Engr.*, **3** (2006), 635–660.

[6] H.T. BANKS AND D. M. BORTZ, A parameter sensitivity methodology in the context of HIV delay equation models, *CRSC-TR02-24*, NCSU, August, 2002; *J. Mathematical Biology*, **50** (2005), 607–625.

[7] H.T. BANKS, D.M. BORTZ AND S.E. HOLTE, Incorporation of variability into the mathematical modeling of viral delays in HIV infection dynamics, *Mathematical Biosciences*, **183** (2003), 63–91.

[8] H.T. BANKS, D.M. BORTZ, G.A. PINTER AND L.K. POTTER, Modeling and imaging techniques with potential for application in bioterrorism, Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, (H.T. Banks and C. Castillo-Chavez, eds.), Frontiers in Applied Mathematics **FR28**, SIAM, Philadelphia, 2003, pp. 129–154.

[9] H.T. BANKS, M. BUKSAS AND T. LIN, *Electromagnetic Material Interrogation Using Conductive Interfaces and Acoustic Wavefronts*, Frontiers in Applied Mathematics **FR21**, SIAM, Philidelphia, 2000.

[10] H.T. BANKS AND J. A. BURNS, Hereditary control problems: Numerical methods based on averaging approximations , *SIAM J. Control Optimization*, **16** (1978), 169–208.

[11] H.T. BANKS AND N.L. GIBSON, Well-posedness in Maxwell systems with distributions of polarization relaxation parameters, *CRSC-TR04-01*, NCSU, January, 2004; *Applied Math. Letters*, **18** (2005), 423–430.

[12] H.T. BANKS AND N.L. GIBSON, Electromagnetic inverse problems involving distributions of dielectric mechanisms and parameters (with N.L. Gibson), *CRSC-TR05-29*, NCSU, August, 2005; *Quarterly of Applied Mathematics*, to appear.

[13] H.T. BANKS AND F. KAPPEL, Spline approximations for functional differential equations, *J. Differential Equations*, **34** (1979), 496–522.

[14] H.T. BANKS AND K. KUNISCH, *Estimation Techniques for Distributed Parameter Systems*, Boston:Birkhaüser, Boston, 1989.

[15] H.T. BANKS, A.J. KURDILA AND G. WEBB, Identification of hysteretic control influence operators representing smart actuators, part I: Formulation, *Mathematical Problems in Engineering*, **3** (1997), 287–328.

[16] H.T. BANKS, A.J. KURDILA AND G. WEBB, Identification of hysteretic control influence operators, part II: Convergent approximations, *J. Intelligent Material Systems and Structures*, **8** (1997), 536–550.

[17] H.T. BANKS, N.G. MEDHIN AND G.A. PINTER, Multiscale considerations in modeling of nonlinear elastomers, *CRSC-TR03-42*, NCSU, October, 2003; *Intl. J. Comp. Meth. Engr. Sci. and Mech.*, to appear.

[18] H.T. BANKS, N.G. MEDHIN AND G.A. PINTER, Nonlinear reptation in molecular based hysteresis models for polymers, *CRSC-TR03-45*, NCSU, December, 2003; *Quart. Appl. Math.*, **62** (2004), 767–779.

[19] H.T. BANKS AND H. K. NGUYEN, Sensitivity of dynamical systems to Banach space parameters, *CRSC-TR05-13*, NCSU, February, 2005; *Journal of Mathematical Analysis and Applications*, **323** (2006), 146–161.

[20] H.T. BANKS AND G.A. PINTER, A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue, *CRSC-TR04-03*, NCSU, January, 2004; *SIAM J. Multiscale Modeling and Simulation*, **3** (2005), 395–412.

[21] H.T. BANKS, R.C. SMITH AND Y. WANG, *Smart Material Structures: Modeling, Estimation and Control*, Masson/John Wiley, Paris/Chichester, 1996.

[22] J.J. BATZEL, F. KAPPEL, D. SCHNEDITZ AND H.T. TRAN, *Cardiovascular and Respiratory Systems: Modeling, Estimation and Control*, Frontiers in Applied Mathematics, SIAM, Philadelphia, 2006.

[23] P. BILLINGSLEY, *Convergence of Probability Measures*, Wiley, NY, 1968.

[24] S.M. BLOWER AND H. DOWLATABADI, Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model as an example, *International Statistics Review*, **62** (1994), 229–243.

[25] J. BORGGAARD AND J.A. BURNS, A PDE sensitivity equation method for optimal aerodynamic design, *Journal of Computational Physics*, **136** (1997), 367–384.

[26] J. Borggaard and A. Verma, On efficient solutions to the continuous sensitivity equation using automatic differentiation, *SIAM Journal on Scientific Computing*, **22** (2001), 39–62.

[27] C. J. F. Böttcher and P. Bordewijk, *Theory of Electric Polarization*,Vol. II, Elsevier, New York, 1978.

[28] M. Brokate and J. Sprekels, *Hysteresis and Phase Transitions*, Springer-Verlag, New York, NY, 1996.

[29] K.P. Burnham and D.R. Anderson, *Model Selection and Multimodal Inference: A Practical Information-Theoretic Approach*, Springer-Verlag, New York, NY, 2002.

[30] J.A. Burns, T. Lin and L. Stanley, A Petrov-Galerkin finite element method for interface problems arising in sensitivity computations, *Computers and Mathematics with Applications*, **49** (2005), 1889–1903.

[31] J. A. Burns, D. Rubio and M. I. Troparevsky, Sensitivity computations for elliptic equations with interfaces, in *Proceedings ICNPAA-2006 Conference on Mathematical Problems in Engineering and Aerospace Sciences*, to appear.

[32] J.A. Burns and L. Stanley, A note on the use of transformations in sensitivity computations for elliptic systems, *Mathematical and Computer Modelling*, **33** (2001), 101–114.

[33] E. M. Cliff, M. Heinkenschloss and A.R. Shenoy, Adjoint-based methods in aerodynamic design-optimization, in *Computational Methods in Optimal Design and Control*, J.T. Borggaard, J. Burns, E. Cliff, and S. Schreck, eds, Birkhauser-Boston, 1998, 91–112.

[34] J.B. Cruz, ed., *System Sensitivity Analysis*, Dowden, Hutchinson & Ross, Inc., Stroudsberg, PA, 1973.

[35] M. Davidian and D.M. Gilitan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1995.

[36] M. Eslami, *Theory of Sensitivity in Dynamic Systems: An Introduction*, Springer-Verlag, Berlin, 1994.

[37] K.R. Foster and H.P. Schwan, Dielectric properties of tissues and biological materials: A critical review, *Critical Rev. in Biomed. Engr.* **17** (1989), 25–104.

[38] P.M. Frank, *Introduction to System Sensitivity Theory*, Academic Press, Inc., New York, NY, 1978.

[39] K.W. Hasson, D.V Lightner, L.L. Mohney, R.M. Redman, B.T. Poulos and B.M. White, Taura Syndrome Virus (TSV) lesion development and the disease cycle in the Pacific white shrimp *Penaeus vannamei*, *Diseases of Aquatic Organisms*, **36** (1999), 81–93.

[40] E.J. Haug, K.K. Choi and V. Komkov, *Design Sensitivity Analysis of Structural Systems*, Academic Press, New York, NY, 1986.

[41] M. Kleiber, H. Antunez, T.D. Hien and P. Kowalczyk, *Parameter Sensitivity in Nonlinear Mechanics: Theory and Finite Element Computations*, John Wiley & Sons, New York, NY, 1997.

[42] Mark Kot, *Elements of Mathematical Ecology*, Cambridge University Press, Cambridge, 2001.

[43] M. Krasnoselskii and A. Pokrovskii, *Systems with Hysteresis*, Nauka/Springer-Verlag, 1983/1989.

[44] E.L. Lehmann, *Elements of Large-Sample Theory*, Springer-Verlag, New York, NY, 1999.

[45] I. Mayergoyz, *Mathematical Models of Hysteresis*, Springer-Verlag, New York, NY, 1991.

[46] E.J. McShane, Generalized curves, *Duke Math J.*, **6** (1940), 513–536.

[47] E.J. McShane, Relaxed controls and variational problems, *SIAM J. Control*, **5** (1967), 438–484.

[48] J.A.J. Metz and E.O. Diekmann, *The Dynamics of Physiologically Structured Populations*, Lecture Notes in Biomathematics, Vol. **68**, Springer, Heidelberg, 1986.

[49] *The NCSU Entomology Homepage,* on line at http://ipm.ncsu.edu/AG271/forages/pea_aphid.html.

[50] A. Saltelli, K. Chan and E.M. Scott, eds., *Sensitivity Analysis*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, 2000.

[51] W.W.Schmaedeke, Optimal control theory for nonlinear vector differential equations containing measures, *SIAM J. Control*, **3** (1965), 231–280.

[52] G.A.F. Seber and C.J. Wild, *Nonlinear Regression*, John Wiley & Sons, Inc., New York, NY, 1989.

[53] L.B. Sheiner, B. Rosenberg and K. Melmon, Modeling of individual pharmacokinetics for computer-aided drug dosage, *Comp. Biomed. Res.*, **5** (1972), 441–459.

[54] L. Stanley and D. Stewart, *Design Sensitivity Analysis: Computational Issues of Sensitivity Equation Methods*, Frontiers in Applied Mathematics, **FR25**, SIAM, Philidelphia, 2002.

[55] J. D. Stark and T. M. Rangus, Lethal and sublethal effects of the neem insecticide formulation, 'Margosan-O', on the pea aphid, *Pestic. Sci.,* **41** (1994), 155-160.

[56] E.R Von Schweidler, Studien über anomalien im verhalten der dielektrika, *Ann. Physik*, **24** (1907), 711–770.

[57] K.W. Wagner, Zur theorie der unvollkommenen dielektrika, *Ann. Physik*, **40** (1913), 817–855.

[58] J. Warga, Relaxed variational problems, *J. Math. Anal. Appl.*, **4** (1962), 111–128.

[59] J. Warga, Functions of relaxed controls, *SIAM J. Control*, **5** (1967), 628–641.

[60] J. Warga, *Optimal Control of Differential and Functional Equations*, Acadedmic Press, New York, NY, 1972.

[61] A. Wierzbicki, *Models and Sensitivity of Control Systems*, Studies in Automation and Control **5**, Elsevier Science Publ. Co., Inc., New York, NY, 1984.

[62] L.C. Young, Generalized curves and the existence of an attained absolute minimum in the calculus of variations,*C. R. Soc. Sci. et Lettres, Varsovie, Cl. III*, **30** (1937), 212–234.

[63] L.C. Young, Necessary conditions in the calculus of variations, *Acta Math.*, **69** (1938), 239–258.