2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Delta-Method Induced Confidence Bands for a Parameter-Dependent Evolution System with Application to Transdermal Alcohol Concentration Monitoring*

Haoxing Liu[1], Larry Goldstein[1], Susan E. Luczak[2] and I. G. Rosen[1]

***Abstract*— Uncertainty caused by parameter randomness in a system modeling the relationship between alcohol concentration level in the blood (BAC) or breath (BrAC) and transdermal alcohol concentration (TAC) measured on the surface of the skin by a wearable, non-invasive, electro-chemical biosensor is considered. The parameter-dependent impulse response function (IRF) and system output in the form of a convolution are expressed in terms of an analytic semigroup of operators with regularly dissipative generator set in a Gelfand triple of Hilbert spaces. The Fréchet derivative of the analytic semigroup in this setting is used to study the variation in the response function resulting from the uncertainties in the parameters described by probability distributions whose statistics depend on regression models with covariates as predictor variables. Finite dimensional approximation of the infinite dimensional state space system and the multi-variate delta method for nonlinear functions of random vectors with asymptotically normal distributions are used to obtain approximating uniform confidence bands for the IRF and the TAC output signal. Convergence of the approximations and three different techniques for obtaining the confidence bands are analyzed and compared.**

I. INTRODUCTION.

We consider a SISO abstract parabolic population model for the transdermal transport of ethanol from the blood-rich dermal layer of the skin through the blood-poor epidermal layer and its collection and measurement by an electro-chemical biosensor on the skin surface. The model takes the form a hybrid ODE/PDE initial boundary value problem with two unknown and directly unmeasurable parameters which vary with covariates that include characteristics of the study participant or patient wearing the sensor, environmental conditions (e.g ambient temperature and humidity) and the manufacturer and bench calibration of the senor hardware. The input to the model is blood or breath alcohol concentration (BAC/BrAC) and the output is transdermal alcohol concentration or level (TAC). The ultimate goal is to deconvolve a BAC/BrAC estimate from TAC observations.

Our model is a population model in that we consider the unknown parameters to be the response variables of a statistical regression with the covariates as predictor variables. In our study here we focus on quantifying the propagation of the uncertainty in the parameters through the model to the impulse response function (IRF) and the convolved TAC output signal. Our approach is based on a Taylor series-based statistical technique known as the multi-variate delta method [2], [3] and the Fréchet derivative of the analytic semigroup [13] that describes the evolution of the system. In an earlier effort [6] we derived conservative naive confidence bands; here we construct tighter uniform confidence bands via techniques specifically applicable to convolutions [8].

*This research supported in part under NIH (NIAAA) grants no. R21AA017711 and R01AA026368.

[1]Haoxing Liu, Larry Goldstein, and I.G. Rosen are with the Department of Mathematics, University of Southern California, Los Angeles, CA USA 90089-2532 haoxingl@usc.edu, larry@usc.edu and grosen@math.usc.edu

[2] Susan E. Luczak is with the Department of Psychology, University of Southern California, Los Angeles, CA USA 90089 luczak@usc.edu

II. THE EVOLUTION SYSTEM AND ITS FINITE DIMENSIONAL APPROXIMATION.

In this section We summarize the underlying model, its finite dimensional approximation, and its input/output map in the form of a convolution. Details can be found in [4], [6], [7].

A. A First Principles Model for Transdermal Transport of Ethanol

We consider a first principles, linear, hybrid, SISO system for the transdermal transport of ethanol from the blood to the sensor. For $t > 0$ and $0 < \eta < 1$, we have

$$\begin{aligned} \frac{\partial x}{\partial t}(t,\eta) &= q_1 \frac{\partial^2 x}{\partial \eta^2}(t,\eta), \\ \frac{dw}{dt}(t) &= q_1 \frac{\partial x}{\partial \eta}(t,0), \quad y(t) = w(t), \end{aligned} \tag{1}$$

with $x(t,0) = w(t)$, $x(t,1) = q_2 u(t)$, $w(0) = w_0$, and $x(0,\eta) = \varphi_0(\eta)$. The ODE describes the inflow/outflow dynamics of the sensor itself, while the parabolic PDE is intended to capture the diffusion of ethanol molecules within the interstitial fluid between the cells in the epidermal layer of the skin. The epidermal layer does not have an active blood supply; its constituent components are made up of a combination of living and dead cells that obtain nourishment from the dermal layer of the skin via the interstitial fluid. In (1), $x(t,\eta)$ is the concentration of ethanol in the epidermal layer at time t and dimensionless depth $\eta \in [0,1]$. The ethanol concentration in the biosensor collection chamber is $w(t)$. The output is the observed TAC signal, $y(t)$, and it is equal to $w(t)$. The input, the BAC or BrAC signal, is $u(t)$. There are two unmeasurable, physiologically-, hardware-, and environmentally-dependent parameters, $q = [q_1, q_2]$. The dimensionless q_1 is essentially the diffusivity of ethanol in the epidermal interstitial fluid and the dimensionless q_2

models the product of the alcohol impedance between the epidermal and dermal layers and the alcohol impedance between the epidermal layer and the membrane covering the collection chamber of the sensor where it contacts the skin.

Note that the input to (1) is on the boundary of the spatial domain $[0,1]$. However, the sensor has a sampling time $\tau > 0$ and the sampled time state space formulation has the advantage that the input operator is bounded [4]. We assume zero-order hold input, $u(t) = u_k, t \in [k\tau, (k+1)\tau), k = 0,1,2,\ldots,K$ where $T = K\tau$. We set $x_k = x_k(\eta) = x(k\tau,\eta), w_k = w(k\tau)$, and $y_k = y(k\tau), k = 0,1,\ldots,K$. Letting $v(t,\eta) = x(t,\eta) - \xi(\eta)u_k$ with $\xi(\eta) = q_2\eta$, for $0 < \eta < 1$, and $k\tau \le t < (k+1)\tau$, each $k = 0,1,2,\ldots,K$, (1) becomes $\frac{\partial v}{\partial t}(t,\eta) = q_1\frac{\partial^2 v}{\partial \eta^2}(t,\eta)$, $\frac{dw}{dt}(t) = q_1\frac{\partial v}{\partial \eta}(t,0) + q_2u_k$, with $v(t,0) = w(t)$, $v(t,1) = 0$, and $v(k\tau,\cdot) = x_k - \xi u_k$.

B. The State Space Formulation

Let $\mathscr{Q}$ be a compact subset of the positive orthant of $\mathbb{R}^2$. Define $H = \mathbb{R} \times L^2(0,1)$ with the inner product $\langle(\theta_1,\varphi_1),(\theta_2,\varphi_2)\rangle = \theta_1\theta_2 + \int_0^1 \varphi_1(\eta)\varphi_2(\eta)d\eta$. Further, let $V = \{(\theta,\varphi) \in H : \varphi \in H^1(0,1), \theta = \varphi(0), \varphi(1) = 0\}$ be a Hilbert space with inner product $\langle(\varphi_1(0),\varphi_1),(\varphi_2(0),\varphi_2)\rangle_V = \langle\varphi_1',\varphi_2'\rangle_{L^2(0,1)}$, where $\langle\cdot,\cdot\rangle_{L^2(0,1)}$ denotes the standard $L^2(0,1)$ inner product. Let $|\cdot|$ and $||\cdot||$ denote respectively the norms induced by these two inner products on their respective spaces. In this way, standard arguments yield the dense and continuous embeddings $V \hookrightarrow H \hookrightarrow V^*$ [14].

For each $q \in \mathscr{Q}$, $a(q;\cdot,\cdot) : V \times V \to \mathbb{R}$ is $a(q;\hat{\varphi}_1,\hat{\varphi}_2) = q_1\int_0^1 \varphi_1'(\eta)\varphi_2'(\eta)d\eta$, $\varphi_1,\varphi_2 \in V$ where $\hat{\varphi}_1 = (\varphi_1(0),\varphi_1), \hat{\varphi}_2 = (\varphi_2(0),\varphi_2)$. It is easily shown that this sesquilinear form is bounded and coercive uniformly in q for $q \in \mathscr{Q}$ and that it is continuous and affine with respect to $q \in \mathscr{Q}$ [4]. Define the operator $A(q) : \text{Dom}(A(q)) \subset H \to H$ by letting $\langle A(q)\hat{\varphi},\hat{\psi}\rangle_{V^*,V} = -a(q;\hat{\varphi},\hat{\psi})$ and setting $\text{Dom}(A(q)) = \{\hat{\phi} = (\varphi(0),\varphi) \in V : \varphi \in H^2(0,1)\}$. Note that the domain $\text{Dom}(A(q))$ does not depend on q, and for $\hat{\varphi} \in \text{Dom}(A(q))$, we have $A(q)\hat{\varphi} = A(q)(\varphi(0),\varphi) = (q_1\varphi'(0), q_1\varphi'')$. Standard arguments yield $A(q)$ densely defined on H, regularly dissipative and self-adjoint [1], [16]. Hence, $A(q)$ is the infinitesimal generator of a uniformly exponentially stable, self-adjoint, analytic semigroup of bounded linear operators, $\{T(q;t) : t \ge 0\}$, or $\{e^{A(q)t} : t \ge 0\}$, on H [1], [11], [16].

With these definitions, the model on $k\tau \le t < (k+1)\tau$ can be rewritten as $\frac{d\hat{v}}{dt}(t) = A(q)\hat{v}(t) + q_1q_2\,(1,0)\,u_k$, with initial conditions $\hat{v}(k\tau) = (w_k, x_k - \xi u_k)$. We set $\hat{x_k} = (w_k,x_k)$, and let $\hat{A}(q) = T(q;\tau) = e^{A(q)\tau} \in \mathscr{L}(H,H)$, $B(q) = q_1q_2\,(1,0) \in \mathscr{L}(\mathbb{R}^1,H)$, and $\hat{B}(q) \in \mathscr{L}(\mathbb{R}^1,H)$ by $\hat{B}(q) = \left(I - \hat{A}(q)\right)((0,\xi) - \int_0^\tau e^{A(q)s})B(q)ds = q_2\left(I - \hat{A}(q)\right)((0,\eta) - q_1A(q)^{-1}(1,0)) \in \mathscr{L}(\mathbb{R}^1,H)$. Note that $A(q)$ elliptic implies $A(q)^{-1} \in \mathscr{L}(H,H)$ exists.

Recalling that $v(t,\eta) = x(t,\eta) - \xi(\eta)u_k$, these definitions together with the variation of constants formula yield the state space form of the model as $\hat{x}_{k+1} = (w_{k+1},x_{k+1}) = (w((k+1)\tau), x((k+1)\tau,\cdot)) = \hat{v}((k+1)\tau) + (0,\xi)u_k = e^{A(q)\tau}(w_k, x_k - \xi u_k) + q_1q_2\int_0^\tau e^{A(q)s}\,(1,0)\,ds u_k + (0,\xi)u_k$. Setting the output operator $\hat{C} \in \mathscr{L}(H,\mathbb{R})$ as $\hat{C}(\theta,\phi) = \theta$ for $(\theta,\phi) \in H$, the discrete time model now becomes $\hat{x}_{k+1} = \hat{A}(q)\hat{x}_k + \hat{B}(q)u_k$, $y_k = \hat{C}\hat{x}_k$, $k = 0,1,2,\ldots,K$. We assume that neither the epidermal layer nor the sensor collection chamber contains any alcohol at time $t = 0$, so $\hat{x}_0 = (w_0,x_0) = (0,0)$. Consequently, the output y of the discrete time system can then be written as a discrete time convolution of the input, u, with a filter, $h(q)$, as $y_k = \sum_{j=0}^{k-1}\hat{C}\hat{A}(q)^{k-j-1}\hat{B}(q)u_j = \sum_{j=0}^{k-1}h_{k-j-1}(q)u_j$, $k = 1,2,...,K$. The convolution kernel or filter is given by $h_i(q) = \hat{C}\hat{A}(q)^i\hat{B}(q), i = 0,1,2,...,K-1$.

For $N = 1,2,...$, let $\{\varphi_j^N\}_{j=0}^N \subset H^1(0,1)$ be the set of linear splines on the interval $[0,1]$ corresponding to the uniform mesh $\{0,\frac{1}{N},\frac{2}{N},...,1\}$ given by $\varphi_0^N(x) = 1 - Nx$ if $x \in [0,\frac{1}{N}]$, $\varphi_0^N(x) = 0$ otherwise, $\varphi_N^N(x) = Nx - N + 1$ if $x \in [\frac{N-1}{N},1]$, $\varphi_N^N(x) = 0$ otherwise, and for $j = 1,2,...,N-1$, $\varphi_j^N(x) = Nx - j + 1$ if $x \in [\frac{j-1}{N},\frac{j}{N}]$, $\varphi_j^N(x) = j+1-Nx$ if $x \in [\frac{j}{N},\frac{j+1}{N}]$, $\varphi_j^N(x) = 0$ otherwise. Let $V^N = \text{span}\{\hat{\varphi}_j^N\}_{j=0}^{N-1} = \text{span}\{(\varphi_j^N(0),\varphi_j^N)\}_{j=0}^{N-1}$. Then V^N is a subspace of V, and the properties of splines [1], [12] yield that for any $v \in V$, there exists a $u^N \in V^N$ with $||u^N - v||_V \to 0$ as $N \to \infty$.

Let $P^N : H \to V^N$ be the orthogonal projection of H onto V^N and let $A^N(q) \in \mathscr{L}(V^N,V^N)$ be given by $< A^N(q)\hat{\varphi}^N,\hat{\psi}^N >= -a(q;\hat{\varphi}^N,\hat{\psi}^N)$, for $\hat{\varphi}^N,\hat{\psi}^N \in V^N$. It follows that P^N converges strongly to the identity on H. We set $\hat{A}^N(q) = e^{A^N(q)\tau} \in \mathscr{L}(V^N,V^N)$, $\hat{B}^N(q) = q_2(I - \hat{A}^N(q))\left(P^N(0,\eta) - q_1A^N(q)^{-1}P^N\,(1,0)\right) \in \mathscr{L}(\mathbb{R}^m,V^N)$, and $\hat{C}^N = \hat{C} \in \mathscr{L}(V^N,\mathbb{R})$. The definition of $A^N(q)$, the compactness of $\mathscr{Q}$, and the uniform coercivity of $a(q;\cdot,\cdot)$ yield the uniform exponential bound on $e^{A^N(q)t}$ and that $R_\lambda(A^N(q))P^N\hat{\varphi} \to R_\lambda(A(q))\hat{\varphi}$ as $N \to \infty$ in H, V and V^* for every $\hat{\varphi} \in H$ (or V or V^*, as the case may be). It follows that $\hat{A}^N(q) = e^{A^N(q)\tau}P^N$ converges strongly to $\hat{A}(q) = T(q;\tau)$ on H uniformly in q for $q \in \mathscr{Q}$.

The approximating model is then given by $y_k^N = \sum_{j=0}^{k-1}h_{k-j-1}^N(q)u_j$, $k = 1,2,...K$, where $h_i^N(q) = C^N\hat{A}^N(q)^i\hat{B}^N(q)$, $i = 0,1,2,...K-1$, with $h_i^N(q)$ converging to $h_i(q)$ in $\mathbb{R}$, uniformly in q for $q \in \mathscr{Q}$ and i in $\{0,1,2,...,K-1\}$ and y_k^N converging to y_k in $\mathbb{R}$, uniformly in q for $q \in \mathscr{Q}$ and k in $\{1,2,...,K\}$.

C. The Fréchet Derivative of the Analytic Semigroup.

The delta method requires the differentiation of the IRF which in turn requires the Fréchet derivative of the analytic semigroup, $DT(q_0;t)$, at $q_0 \in \mathscr{Q}$. We consider the sector in $\mathbb{C}$ contained in the resolvent set $\rho(A(q))$ for all $q \in \mathscr{Q}$ given by

$$\Sigma_\gamma = \left\{\lambda \in \mathbb{C} |\arg(\lambda - \lambda_0)| \le \frac{\pi}{2} + \theta_\gamma\right\}$$

where $\gamma \in (0,1)$ is a constant, and $\alpha_1 = 1 + \alpha_0/\mu_0$ and $\theta_\gamma = \tan^{-1}\left(\alpha_1/(1-\gamma)\right)$ where α_0, μ_0, and λ_0 are the (uniform in $q \in \mathscr{Q}$) boundedness and coercivity constants (see [6]) for the form $a(q;\cdot,\cdot) : V \times V \to \mathbb{R}$. It can then be shown [11], [16] that the resolvent of the infinitesimal generator

$A(q)$, $R(\lambda;q) = \big(\lambda I - A(q)\big)^{-1}$, is uniformly bounded for all $\lambda \in \Sigma_\gamma$ in the sense that $\|R(\lambda,q)\|_{V^*,V} \le 1/(\alpha_2\gamma)$ [13]. It then follows that the map $q \to R(\lambda,q)$ has a uniformly (in Σ_γ) norm-convergent power series representation in terms of $A(\delta q)$ and thus has Fréchet derivative $\frac{D}{Dq}R(\lambda,q)$ acting on δq as a linear map from $\mathscr{Q}$ to $\mathscr{L}(V^*,V)$. Moreover, for fixed $\gamma \in (0,1)$ and $q \in \mathscr{Q}$, $\frac{D}{Dq}R(\lambda,q)\delta q = R(\lambda,q)A(\delta q)R(\lambda,q)$ [13].

In addition, the operator $T(q;t)$ can be expressed as an operator contour integral of the form $T(q;t) = \frac{1}{2\pi i}\int_{\partial\Sigma_\gamma} e^{\lambda t}R(\lambda,q)\,\mathrm{d}\lambda$ for $t > 0$ and $q \in \mathscr{Q}$ [11]. Combining this with the uniform boundedness and analyticity of $R(\lambda,q)$, and the dominated convergence theorem, one can conclude [13] that the Fréchet derivative of operator $T(q;t)$ acting on δq exists and is given by

$$DT(q;t)\delta q = \frac{1}{2\pi i}\int_{\partial\Sigma_\gamma} e^{\lambda t}R(\lambda,q)A(\delta q)R(\lambda,q)\mathrm{d}\lambda. \tag{2}$$

Analogously, from the definition of the approximating semigroup, particularly the definition of its infinitesimal generator, the coercivity and boundedness conditions must be satisfied for all N. Thus we have that Σ_γ, which depends only on the coercivity and boundedness constants, is contained in the resolvent set $\rho(A^N(q))$ and that for $q \in \mathscr{Q}$, the resolvent operator $R^N(\lambda,q)$ is uniformly bounded for all $\lambda \in \Sigma_\gamma$. The operators $T^N(q;t)$ can be expressed as a contour integral of the form $T^N(q;t) = \frac{1}{2\pi i}\int_{\partial\Sigma_\gamma} e^{\lambda t}R^N(\lambda,q)\,\mathrm{d}\lambda$ for $t > 0$ and $q \in \mathscr{Q}$ [11]. We conclude that $\frac{D}{Dq}R^N(\lambda,q)\delta q = R^N(\lambda,q)A^N(\delta q)R^N(\lambda,q)$. The dominated convergence theorem and uniform boundedness of the resolvent operators then yield

$$\frac{D}{Dq}T^N(q;t)\delta q = \frac{1}{2\pi i}\int_{\partial\Sigma_\gamma} e^{\lambda t}R^N(\lambda,q)A^N(\delta q)R^N(\lambda,q)\,\mathrm{d}\lambda. \tag{3}$$

Under these same assumptions, it can be argued that $\frac{D}{Dq}T^N(q;t)P^N\varphi \to \frac{D}{Dq}T(q;t)\varphi$ in the V^*-norm as $N \to \infty$, uniformly in t for $t > 0$ for every $\varphi \in H$ [6]. The argument is based on the bounds for the resolvent given in Lemma 3.6.1 in [16], and, considering the fact that $A(\delta q) \in \mathscr{L}(V,V^*)$, we have that $R(\lambda,q)A(\delta q)R(\lambda,q) \in \mathscr{L}(H,V^*)$, with $\|R(\lambda,q)A(\delta q)R(\lambda,q)\|_{\mathscr{L}(H,V^*)} \le \frac{L_1}{|\lambda|^{\frac{3}{2}}}$, for some $L_1 > 0$. Analogously we also have that $\|R^N(\lambda,q)A^N(\delta q)R^N(\lambda,q)P^N\|_{\mathscr{L}(H,V^*)} \le \frac{L_1}{|\lambda|^{\frac{3}{2}}}$. Using the strong convergence of the resolvents, $R^N(\lambda,t)P^N \to R(\lambda,t)$ in V and H and density arguments in V^*, the dominated convergence theorem yields the desired result. Finally, the fact that the parameter space, $\mathscr{Q} \subset \mathbb{R}^2$ is finite dimensional, we get that $\frac{D}{Dq}T^N(q;t)P^N\varphi \to \frac{D}{Dq}T(q;t)\varphi$ in $\mathscr{L}(\mathbb{R}^2,V^*)$ as $N \to \infty$ uniformly in t for $t > 0$.

III. The Delta Method.

A. The Multivariate Delta Method

The delta method is based on the following result [2], [3]. If $g: \mathbb{R}^r \mapsto \mathbb{R}^d$ is such that $Dg(x)$ is continuous in a neighborhood of $\mu \in \mathbb{R}^d$, and $\mathbf{X_n}$ is a sequence of r-dimensional random variables such that $\sqrt{n}\,(\mathbf{X_n} - \mu) \to \mathbf{X}$ in distribution for some r-dimensional random variable $\mathbf{X}$, then $\sqrt{n}\,\big(g(\mathbf{X_n}) - g(\mu)\big) \to Dg(\mu)\mathbf{X}$ in distribution as $n \to \infty$

Let $q_0 \in \mathscr{Q} \subseteq \mathbb{R}^r$ be fixed and $\Gamma \in \mathbb{R}^{r\times r}$ be a positive definite matrix. We assume there exists a sequence of random variables $\{\mathbf{q}_n\}$ with support in $\mathscr{Q}$ with constant order expectation $\mathbb{E}[\mathbf{q}_i] = q_0 + o\left(1/\sqrt{n}\right)$ for $i = 1,2...$ such that $\sqrt{n}\,(\mathbf{q_n} - q_0) \to \mathbf{N}(0,\Gamma)$ in distribution as $n \to \infty$ where $\mathbf{N}(0,\Gamma)$ denotes the r-dimensional multivariate normal random variable with mean 0 and covariance matrix Γ. For our study, since the random sequence $\mathbf{q_n}$ usually originates from estimations based on standard regression approaches, the asymptotic normality assumption can be realized with minor conditions on the data.

Now for a multivariate scalar-valued delta method, consider a given function $g: \mathbb{R}^r \mapsto \mathbb{R}$ with continuous first-degree partial derivatives $\frac{\partial g}{q_j}$, $j = 1,...,r$. The gradient of g as $Dg(q) := \left[\frac{\partial g}{q_j}\right]_j \in \mathbb{R}^r$. If $Dg(q_0)$ is non-zero and $\sum_{i,j}\Gamma_{i,j}\frac{\partial g(q_0)}{q_i}\frac{\partial g(q_0)}{q_j} > 0$, then $\sqrt{n}\,\big(g(\mathbf{q_n}) - g(q_0)\big) \to \mathbf{N}\left(0, (Dg(q_0))\Gamma(Dg(q_0))^\intercal\right)$ in distribution as $n \to \infty$.

The purpose of the additional condition on the sum is to ensure that the resulting variance $\Gamma_g := (Dg(q_0))\Gamma(Dg(q_0))^\intercal = \sum_{i,j}\Gamma_{i,j}\frac{\partial g(q_0)}{q_i}\frac{\partial g(q_0)}{q_j}$ is positive, corresponding to the positive definite case for the multivariate normal distribution. In the particular example of interest to us here, the dimension of the parameter space $\mathbb{R}^r$ is typically $r = 1$ or $r = 2$, and the model is SISO and consequently $d = 1$.

B. Delta Method Based Confidence Bands.

It is clear that the operators $\hat{B}(q)$ and $\hat{C}(q)$ are Fréchet differentiable with respect to q. Recalling that $h_k(q) = \hat{C}(q)\hat{A}(q)^k\hat{B}(q) \in \mathbb{R}$, $k = 0,1,...K-1$, we consider the response to a unit impulse at a specific discrete time k which corresponds to $h_k(q)$. Note that $h_k(q): \mathscr{Q} \mapsto \mathbb{R}$ can be viewed as a multivariate scalar-valued function with directional derivatives in the classical sense. In [6] we have shown that $h_k(q) = \left\langle \hat{C}(q), \hat{A}^k(q)\hat{B}(q)\right\rangle_H = \left\langle \hat{C}(q), T(q;k\tau)\hat{B}(q)\right\rangle_H$, where $\hat{B}(q), \hat{C}(q) \in H$ and we can recover the partial derivatives $\frac{\partial h_k(q)}{\partial q_s}$.

It then follows from the delta method that the asymptotic distribution of $h_k(q)$ can be obtained. Indeed, assuming that the partial derivatives of $h_k(q)$ are non-zero at q_0 and denoting $\nabla h_k(q) \in \mathbb{R}^2$ as the gradient, then $\sqrt{n}\,\big(h_k(\mathbf{q_n}) - h_k(q_0)\big)$ converges in distribution to $\mathbf{N}\left(0, (\nabla h_k(q_0))^\intercal\Gamma(\nabla h_k(q_0))\right)$ as $n \to \infty$

Denoting $\Gamma_k = \nabla h_k(q_0))^\intercal\Gamma(\nabla h_k(q_0) \in \mathbb{R}$, the $(1-\alpha)\%$

confidence interval of the response is given by

$$\left(h_k(q_0) - z_{\alpha/2}\frac{1}{\sqrt{n_k}}(\Gamma_k)^{\frac{1}{2}}, h_k(q_0) + z_{\alpha/2}\frac{1}{\sqrt{n_k}}(\Gamma_k)^{\frac{1}{2}}\right)$$

where $z_{\alpha/2}$ denotes the $\frac{\alpha}{2}$-th quantile of the standard normal.

Similarly, the asymptotic distribution of the finite-dimensional approximating convolution is analogous and if $\hat{C}(q) \in V$, then $(Dh_k)^N(q)$ converges to $Dh_k(q)$ uniformly in q for $q \in \mathscr{Q}$ [6]. In particular, the convergence holds for q_0. Hence we conclude that $(\Gamma_k)^N$ converges to Γ_k in $\mathbb{R}$, and so does the confidence band.

Recall that $h_k(q) = \left(\sum_{j=0}^{k} h_{k-j}(q)u_j\right) = y_{k+1}$ is the response of the evolution system to a unit impulse input at time $k = 0$, with $u_0 = 1 \in \mathbb{R}$ and $u_k = 0$ for all $k > 0$. We generalize this result and consider $y_k = y_k(q)$ for some fixed $k = 0, 1, ..., K-1$ with any constant input $\{u_j\}_{j=0}^{k-1}$. Since $h_k(q)$ is differentiable for all $k = 0, 1, ..., K-1$, by linearity, it follows that the response function at time k, $y_k = y_k(q)$, can be viewed as a function from $\mathbb{R}^2$ to $\mathbb{R}$ and is continuously differentiable with derivative $Dy_k(q) \in \mathbb{R}^2$ and whose closed-form can be computed explicitly [6]. Applying the delta-method to $y_k(q)$, we find that if $Dy_k(q_0)$ is non-zero, then

$$\sqrt{n}\left(y_k(\mathbf{q_n}) - y_k(q_0)\right) \to \mathbf{N}\left(0, (Dy_k(q_0))\Gamma(Dy_k(q_0))^\intercal\right)$$

in distribution as $n \to \infty$.

Denoting $(Dy_k(q_0))\Gamma(Dy_k(q_0))^\intercal > 0$ as $\Gamma_k \in \mathbb{R}$, then for large n_k, $y_k = y_k(\mathbf{q_{n_k}})$ is approximately normally distributed with mean $y_k(q_0)$ and variance $\frac{1}{n_k}\Gamma_k$. Since $y_k \in \mathbb{R}$, we construct the $(1-\alpha)\%$ confidence interval in a similar manner.

$$\left(y_k(q_0) - z_{\alpha/2}\frac{1}{\sqrt{n_k}}(\Gamma_k)^{\frac{1}{2}}, y_k(q_0) + z_{\alpha/2}\frac{1}{\sqrt{n_k}}(\Gamma_k)^{\frac{1}{2}}\right)$$

Observe that $y_k(q)$ can be viewed as a linear combination of h_l for $l = 0, 1, ..., K-1$. Consequently, the finite-dimensional response function $Dy_k^N(q)$ converges to $Dy_k(q)$ uniformly in q for all $q \in \mathscr{Q}$. In particular $Dy_k^N(q_0)$ converges element-wise to $Dy_k(q_0)$ in $\mathbb{R}^2$ and Γ_k^N converges to Γ_k in $\mathbb{R}$. It follows that for fixed large n_k, $y_k^N(\mathbf{q_{n_k}})$ converges in distribution to $y_k(\mathbf{q_{n_k}})$ as $N \to \infty$. Hence the confidence band for the finite-dimensional approximation converges to that for the infinite-dimensional system.

IV. Uniform Confidence Bands

We will construct in general three different types of confidence bands for the response $y(q) = (y_1, ..., y_K)$ based on the asymptotic distribution for $y_k(q)$ we presented in the previous section. For each method, we also investigate the convergence of the finite-dimensional approximation. The implementation of the methods will be demonstrated in Section V with an example.

A. Joint Confidence Band Based on Individual Elements

Let $[y_k^-, y_k^+]$ be the corresponding $(1-\alpha)\%$ confidence set for y_k. In a naïve attempt, one can construct the uniform/joint confidence band by connecting all individual confidence bands as $\bigotimes_{k=1}^{K}[y_k^{-\alpha}, y_k^{+\alpha}]$. However, the coverage of the resulting band will be substantially lower than the desired level. If we assume that y_k's are independent, then the confidence level of the naive band is $(1-\alpha)^K$. The method we introduce in this section builds on the same idea of combining individual confidence sets of y_k, but instead, adjusts the levels of confidence at each y_k to account for the loss of coverage in the joint confidence band.

The Bonferroni band is a common variation of the naïve band. For $k = 1, ..., K$, consider the $\left(1 - \alpha/(K)\right)$ confidence level. Denote A_k as the event that the confidence set of y_k contains the true curve, then we can find a lower bound for the confidence coverage of the Bonferroni bound by

$$\mathbb{P}\left(\bigcap_{k=1}^{K} A_k\right) = \mathbb{P}\left(\left(\bigcup_{k=1}^{K} A_i^c\right)^c\right) = 1 - \mathbb{P}\left(\bigcup_{k=1}^{K} A_i^c\right)$$
$$\geq 1 - \sum_{k=1}^{K} \mathbb{P}\left(A_i^c\right) = 1 - K*(\alpha/K) = 1 - \alpha.$$

In other words, the coverage is at least $(1-\alpha)$, giving a conservative estimate for the true curve. It has been noted that if the correlation between individual points is significant, the bandwidth might be substantially larger than the desired level [8]. The joint confidence band is given by $\bigotimes_{k=1}^{K}[y_k^{-\alpha/K}, y_k^{+\alpha/K}]$.

An alternative adjustment for the individual confidence level is suggested by Zbyněk ŠidáK in [15]. The proposed marginal coverage level is $(1-\alpha)^{1/K}$, and the corresponding quantile under normal distribution is given by $1 - \frac{1-(1-\alpha)^{1/K}}{2}$. The coverage level was derived under the assumption that y_k's are independent normal random variables, which agrees with our assumption in III-B. In practice, it is noted in [9] that $(1-\alpha)^{1/K}$ and $\left(1-\alpha/(K)\right)$ are fairly similar with small α and large K, and the resulting confidence bands will have very little difference in the width.

Since the confidence bands described in this section are all computed as deterministic functions of Γ_k, it follows directly that the convergence result in III-B implies the convergence of the confidence bands as well.

B. Uniform Elliptical Confidence Band

So far, the confidence bands we have considered did not account for the correlation between individual time points y_k. In the previous section, we indicated that the adjusted marginal confidence levels do not depend on the distribution of y_k's but rather relied on probability inequalities to achieve a desired overall confidence level. In this section we present the method referred to as the Wald's Band in [5] and [8].

Suppose we have obtained that

$$\sqrt{n}(y - y_0) \to \mathbf{N}(0, \Gamma) \text{ in distribution}$$

for $y, y_0 \in \mathbb{R}^K$ and $\Gamma \in \mathbb{R}^{K\times K}$ positive definite. Then the confidence set, Wald's Ellipse, can be defined as

$$\mathscr{W}(1-\alpha) \coloneqq \left\{ y \in \mathbb{R}^K | (y-y_0)^\intercal \Gamma^{-1}(y-y_0) \le \chi^2_{K,1-\alpha} \right\}$$

where $\chi^2_{K,1-\alpha}$ gives the $(1-\alpha)$ quantitle of the χ^2 distribution with K degrees of freedom. The advantage of this construction is that we obtain a set that gives the exact confidence level we want and the joint distribution incorporates all the possible correlation structures between individual points. However, there are several challenges associated with this approach. First, the confidence set in $\mathbb{R}^K$ is not directly visualizable for higher dimensions. To obtain a confidence band in the form of $\bigotimes_{k=1}^K [y_k^{-\alpha}, y_k^{+\alpha}]$, we will study the projection of $\mathscr{W}(1-\alpha)$ onto the coordinates of y as

$$\mathscr{W}_{CI}(1-\alpha) = \bigotimes_{k=1}^{K} [\min_{y\in\mathscr{W}(1-\alpha)} y_k, \max_{y\in\mathscr{W}(1-\alpha)} y_k]$$

which adds an additional layer of complexity to our computation and convergence argument. Moreover, the method requires that Γ is non-degenerating or invertible, which is, in general, not the case in applications where the dimension of the parameter space r (in our case, $r = 2$) is much less than the size of discretization K. Even though the existence of an asymptotically valid joint confidence set with degenerating covariance matrix has been established in [5], construction of such a confidence set relies on approximation methods such as bootstrapping and the power of the confidence set suffer from significant loss during the projection. Recent studies have further demonstrated that the projection set, which includes the Wald ellipse $\mathscr{W}(1-\alpha)$, is overly conservative even compared to the Bonferroni set discussed above [10].

Convergence of the approximating confidence bands follows from the fact that Γ^N converges to Γ element-wise as in III-B. With the additional assumption that Γ^N and Γ are both positve-definite, we conclude that $(\Gamma^N)^{-1}$ converges to Γ^{-1} element-wise, and thus the confidence set $\mathscr{W}^N(1-\alpha)$ converges to $\mathscr{W}(1-\alpha)$ in the sense that $\forall y \in \mathscr{W}(1-\alpha)$, $\exists$ a sequence $\{y_N\}$ such that $y_N \in \mathscr{W}^N(1-\alpha)$ and $y_N \to y$ element-wise in $\mathbb{R}^K$. The convergence implies the corresponding convergence in min/max of y_k and hence in the projection set $\mathscr{W}_{CI}(1-\alpha)$.

C. Uniform Supremum Bands for the Convolution System

Consider the $(1-\alpha)\%$ confidence band of the form $B(c) \coloneqq \bigotimes_{k=1}^K [y_k^{-\alpha}, y_k^{+\alpha}] = \bigotimes_{k=1}^K [y_k - c_\alpha * \sigma_k, y_k + c_\alpha * \sigma_k]$ for some $c_\alpha > 0$ depending on α. It can be shown that the confidence bands in previous sections all fall into this single-parameter family [9] with different critical values c_α. In this section, we present the asymptotically optimal single-parameter confidence band in this setting.

Assuming that $\sqrt{n}(y-\tilde{y}) \to \mathbf{N}(0,\Gamma)$ in distribution, then according to [10], the asymptotic coverage of any single-parameter confidence band can be realized as

$$\mathbb{P}\left(y \in B(c)\right) \to \mathbb{P}\left(\max_{k\in 1,\dots,K} \frac{|y_k - \tilde{y}_k|}{\sqrt{\Gamma_{kk}}} \le c\right)$$

where $\Gamma_{kk} = \sigma_k$ denotes the k-th diagonal element of Γ. We look for critical values of c such that the probability on the right-hand side of the convergence is exactly $(1-\alpha)$. Note that since the y_k's are jointly normal, the cumulative distribution of $\max_{k\in 1,\dots,K} \frac{|y_k-\tilde{y}_k|}{\sqrt{\Gamma_{kk}}}$ can be computed as the joint maximum of correlated normal random variables, which is a function of the covariance matrix Γ as

$$\mathbb{P}\left(\max_{k\in 1,\dots,K} \frac{|y_k - \tilde{y}_k|}{\sqrt{\Gamma_{kk}}} \le c\right) = \mathbb{P}\left(\bigcap_{k=1}^{K} |y_k - \tilde{y}_k| \le c * \sigma_k\right) = F(c;\Gamma).$$

The critical value c_α^{crit} is determined by the $(1-\alpha)\%$ quantile of the random variable with cdf $F(c;\Gamma)$, with the confidence band known as the Sup-t band then given by $\bigotimes_{k=1}^K [y_k - c_\alpha^{crit} * \sigma_k, y_k + c_\alpha^{crit} * \sigma_k]$. By construction, the confidence band has asymptotic coverage of exactly $(1-\alpha)$, and provides a smaller critical value, or equivalently narrower bandwidth, compared to the other types of single-parameter confidence discussed earlier [10].

For fixed values of c, the function $F(c;\Gamma) = F(\Gamma)$ is continuous with respect to Γ, or to all entries of Γ. Thus our convergence results for Γ^N in the finite-dimensional approximation of the system extend naturally to $F(\Gamma^N) \to F(\Gamma)$ and to the convergence of the critical values.

V. An Example and Numerical Studies

We have $r = 2$ and $A(q)\hat{\varphi} = A(q)(\varphi(0), \varphi) = q_1(\varphi'(0), \varphi'') = q_1 A\hat{\varphi}$, where $A : \text{Dom}(A) \subset H \to H$ is given by $A(\varphi(0), \varphi) = (\varphi'(0), \varphi'')$ with $\text{Dom}(A) = \{\hat{\varphi} = (\varphi(0), \varphi) \in V : \varphi \in H^2(0,1)\}$. It follows that $h_k(q) = \left\langle \hat{C}, \hat{A}^k(q)\hat{B}(q)\right\rangle_H = \left\langle (1,0), e^{q_1 A k\tau}\hat{B}(q)\right\rangle_H$, where $\hat{B}(q) = q_2(I - e^{q_1 A\tau})\left((0,\eta) - q_1 A^{-1}(1,0)\right)$. Letting $\mathscr{M}^N$ and $q_1\mathscr{K}^N$ be the $N\times N$ Mass (Gram) and stiffness matrices corresponding to the V^N basis $\{\varphi_j^N\}_{j=0}^{N-1}$ and the operator $A(q)$, $A^N(q) = -q_1(\mathscr{M}^N)^{-1}\mathscr{K}^N$ and $h_k^N(q) = \left\langle [1,\mathbf{0}^T]^T, e^{-q_1(\mathscr{M}^N)^{-1}\mathscr{K}^N k\tau}\hat{B}^N(q)\right\rangle_{\mathbb{R}^N}$, where $\hat{B}^N(q) = q_2(I - e^{-q_1(\mathscr{M}^N)^{-1}\mathscr{K}^N\tau})\left((\mathscr{M}^N)^{-1}\boldsymbol{\eta} + (\mathscr{K}^N)^{-1}[1,\mathbf{0}^T]^T\right)$, with $\mathbf{0}$ being the zero vector in $\mathbb{R}^{N-1}$ and $\boldsymbol{\eta}$ the vector of inner products of $(0,\eta) \in H$ and the basis $\{\hat{\varphi}_j^N\}_{j=0}^{N-1}$.

Denote $\hat{A}^N(q) = \hat{A}^N(q_1) = e^{-q_1(\mathscr{M}^N)^{-1}\mathscr{K}^N\tau}$, then $h_k^N(q) = \left\langle [1,\mathbf{0}^T]^T, \hat{A}^N(q)^k \hat{B}^N(q)\right\rangle_{\mathbb{R}^N}$ and $\hat{B}^N(q) = q_2(I - \hat{A}^N(q)) \cdot \left((\mathscr{M}^N)^{-1}\boldsymbol{\eta} + (\mathscr{K}^N)^{-1}[1,\mathbf{0}^T]^T\right)$. We note that $\frac{\partial \hat{A}^N(q)}{\partial q_1} = -(\mathscr{M}^N)^{-1}\mathscr{K}^N\tau\hat{A}^N(q)$, $\frac{\partial \hat{A}^N(q)}{\partial q_2} = 0$, $\frac{\partial \hat{B}^N(q)}{\partial q_1} = q_2(-\frac{\partial \hat{A}^N(q)}{\partial q_1})\left((\mathscr{M}^N)^{-1}\boldsymbol{\eta} + (\mathscr{K}^N)^{-1}[1,\mathbf{0}^T]^T\right)$, $\frac{\partial \hat{B}^N(q)}{\partial q_2} = (I - \hat{A}^N(q))\left((\mathscr{M}^N)^{-1}\boldsymbol{\eta} + (\mathscr{K}^N)^{-1}[1,\mathbf{0}^T]^T\right)$, $\frac{\partial h_k^N(q)}{\partial q_1} = E_1 + E_2$, and $\frac{\partial h_k^N(q)}{\partial q_2} = \left\langle [1,\mathbf{0}^T]^T, \hat{A}^N(q)^k\hat{B}^N(q)/q_2\right\rangle_{\mathbb{R}^N}$, where $E_1 = \left\langle [1,\mathbf{0}^T]^T, \hat{A}^N(q)^k \frac{\partial \hat{B}^N(q)}{\partial q_1}\right\rangle_{\mathbb{R}^N}$ and $E_2 = \left\langle [1,\mathbf{0}^T]^T, k\hat{A}^N(q)^{k-1}\left(\frac{\partial \hat{A}^N(q)}{\partial q_1}\right)\hat{B}^N(q)\right\rangle_{\mathbb{R}^N}$.

We assumed q_1 and q_2 are independent with $q_i \sim \mathbf{N}(\mu_i, \sigma_i^2)$, $i = 1,2$, where $\mu_1 = 1.25$, $\mu_2 = 1.56$, and

$\sigma_1^2 = \sigma_2^2 = 0.05^2$. We set $T = 5$ and $\tau = 1/60$ hours (one minute or 1/60 Hz). To emphasize different types of confidence bands, we choose a unit impulse as our input (more complicated examples can be found in [7]). Applying the delta method, we generated the naive 95% confidence band plotted in Figure 1. We consider the single-parameter confidence band family mentioned in Section IV, and compute the marginal critical values for individual confidence intervals. For the Wald's band in Section IV-B, we apply the result in [10] that the critical value of the projected Wald's confidence band is given by the corresponding $\sqrt{\chi^2_{K,1-\alpha}}$ statistic. For the sup-t band in Section IV-C, we apply an empirical method to determine the critical value for the maximum of the jointly normal distribution. Note that the relationship between different bands stated in IV holds for our example, where the coverage width of each band in increasing order is given as: Naïve, Sup-t, Bonferroni/ŠidáK, Wald's Projection.

In addition to the plot in Figure 1, we included a chart of absolute sequential differences in Figure 2 for the confidence bandwidth $\text{CB}_k(q;\alpha,N) \in \mathbb{R}^{T/\tau}$, and Abs. Diff.$(N) = \left\|\text{CB}_{k+1}(q;\alpha,N) - \text{CB}_k(q;\alpha,N)\right\|_1, \text{CB}_0 = 0, k = 1,2,...,T/\tau$. Observe that the differences drop quickly towards zero for most of the confidence bandwidth, which indicates the convergence of confidence bands as N increases. For the sup-t band, the convergence is less convincing due to the fact that approximation is required in order to find the critical value. Being an empirical method, unlike the other cases, the critical values for the sup-t band vary with N.

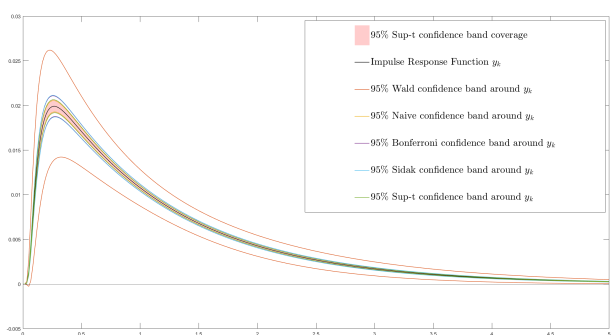

Fig. 1. 95% Confidence Bands for the IRF h_k(black) or output y_k from unit impulse at time $k = 0$ with $N = 60$. The pink region represents the sup-t band coverage, while only upper/lower bounds are provided for the other bands. The continuous band is the interpolation of the discrete confidence intervals discussed in Section IV determined at each time k, $k = 1, 2, ..., K$, where $K = 300$ ($\tau = 1/60$).

VI. Concluding Remarks and Future Research.

Using the delta method, we determined the asymptotic distribution of the response function given the distribution of the parameters q, and established methods for constructing uniform confidence bands. Computations required the introduction of finite-dimensional approximation and the consideration of associated convergence issues. We established convergence of the approximating confidence bands and we presented an example involving a SISO system modeling the transdermal transport of ethanol and numerically illustrated the convergence of the confidence bands for the IRF.

Absolute Sequential Difference for Different Types of Confidence Bands with Increasing N

N	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60
Naïve	3.88E-03	3.52E-04	5.68E-05	1.89E-05	8.63E-06	4.67E-06	2.80E-06	1.82E-06	1.24E-06	8.89E-07	6.57E-07	4.99E-07	3.89E-07	3.08E-07	2.49E-07
Sup-t	4.25E-03	1.83E-03	2.04E-03	1.83E-03	4.60E-04	4.90E-04	1.72E-03	3.51E-03	1.39E-03	2.39E-03	5.73E-03	6.00E-03	2.16E-03	6.02E-03	4.89E-04
Bonferroni	7.10E-03	6.45E-04	1.04E-04	3.46E-05	1.58E-05	8.54E-06	5.13E-06	3.32E-06	2.28E-06	1.63E-06	1.20E-06	9.14E-07	7.11E-07	5.64E-07	4.55E-07
Sidak	7.44E-03	6.75E-04	1.09E-04	3.63E-05	1.66E-05	8.95E-06	5.38E-06	3.48E-06	2.38E-06	1.70E-06	1.26E-06	9.58E-07	7.45E-07	5.91E-07	4.77E-07
Wald	3.69E-02	3.35E-03	5.41E-04	1.80E-04	8.23E-05	4.45E-05	2.67E-05	1.73E-05	1.18E-05	8.47E-06	6.26E-06	4.76E-06	3.70E-06	2.94E-06	2.37E-06

Fig. 2. Table of Absolute Sequential Difference for Confidence Bandwidth

The problem of interest to us here has not been entirely resolved. We are especially interested in finding confidence bands for the inverse or deconvolution problem to quantify the uncertainty in the deconvolved input signal (i.e. the estimated BAC or BrAC) [4] due to the uncertainty in the model parameters. Research on this problem is continuing.

References

[1] H. T. Banks and K. Ito. A unified framework for approximation in inverse problems for distributed parameter systems. *CTAT*, 4(1):73–90, 1988.

[2] George Casella and Roger L Berger. *Statistical Inference*, volume 2. Duxbury, Pacific Grove, CA, 2002.

[3] Thomas S. Ferguson. *A Course in Large Sample Theory (1st Ed.)*. Chapman & Hall/CRC, 1996.

[4] K. Hawekotte, S. E. Luczak, and I. G. Rosen. Deconvolving breath alcohol concentration from biosensor measured transdermal alcohol level under uncertainty: a bayesian approach. *Mathematical Biosciences and Engineering*, 18(5), 2021.

[5] Atsushi Inoue and Lutz Kilian. Joint confidence sets for structural impulse responses. *Journal of Econometrics*, 192(2):421–432, 2016. Innovations in Multiple Time Series Analysis.

[6] Haoxing Liu, L. Goldstein, S. E. Luczak, and I. G. Rosen. Confidence bands for evolution systems described by parameter-dependent analytic semigroups. In *Proceedings of the 2023 SIAM Conference on Control and its Applications*, 2023, to appear.

[7] Haoxing Liu, L. Goldstein, S. E. Luczak, and I. G. Rosen. Delta method confidence bands for convolutions and deconvolutions of evolution systems described by parameter-dependent analytic semigroups with regularly dissipative generators. 2024, in preparation.

[8] Helmut Lütkepohl, Anna Staszewska-Bystrova, and Peter Winker. Constructing joint confidence bands for impulse response functions of var models. *Econometrics and Statistics*, 13:69 – 83, 09 2018.

[9] José Luis Montiel Olea and Mikkel Plagborg-Møller. Simultaneous confidence bands: Theoretical comparisons and recommendations for practice. *Unpublished manuscript*, 2017.

[10] José Luis Montiel Olea and Mikkel Plagborg-Møller. Simultaneous confidence bands: Theory, implementation, and an application to svars. *Journal of Applied Econometrics*, 34(1):1–17, 2019.

[11] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Applied Mathematical Sciences. Springer, 1983.

[12] M. H. Schultz. *Spline Analysis*. Prentice-Hall Series in Automatic Computation. Pearson Education, Limited, 1972.

[13] S Seubert and J.G Wade. Fréchet differentiability of parameter-dependent analytic semigroups. *Journal of Mathematical Analysis and Applications*, 232(1):119 – 137, 1999.

[14] Ralph E Showalter. *Hilbert Space Methods in Partial Differential Equations*. Courier Corporation, 2010.

[15] Zbynek Sidak. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62(318):626–633, 1967.

[16] H. Tanabe. *Equations of Evolution*. Monographs and Studies in Mathematics. Pitman, London, 1979.