

**PROBABILISTIC RECOVERY OF NEUROENDOCRINE PULSATILE,  
SECRETORY AND KINETIC STRUCTURE:  
AN ALTERNATING DISCRETE AND CONTINUOUS SCHEME**

BY

SOMESH CHATTOPADHYAY (*Department of Statistics, Florida State University, Tallahassee, FL 32306-4330*),

DANIEL M. KEENAN (*Department of Statistics, University of Virginia, Charlottesville VA 22904*),

AND

JOHANNES D. VELDHUIS (*Division of Endocrinology and Metabolism, Department of Internal Medicine, Mayo Clinic and Mayo Graduate School of Medicine, Rochester, MN 55905*)

**Abstract.** The brain (hypothalamus) directs hormone secretion by the pituitary gland via burst-like (pulsatile) release of specific peptides at inferentially random times. These pulsatile signals supervise growth, reproduction, lactation, stress adaptations, water balance and immune responses. However, hypothalamic molecules are diluted  $> 3000$ -fold in systemic blood, leaving pituitary-hormone pulses as measurable surrogates. The latter (roughly) mirror hypothalamic peptide bursts on a 1:1 basis, albeit being observed in a noisy environment. As a window to the brain, one must accurately recover the pulse (onset) times, and thereby estimate hormone secretion and kinetic parameters ( $\theta \in \bar{\Theta}$ ) without distortion. Based upon limited observed data, one would like to obtain probability statements about underlying pulsatility, secretion and kinetics. Moreover, to be applicable in today's clinical setting, it is important that any such procedure require minimal or no human input. We propose and justify the following method. First, the data (a pituitary hormone concentration time-profile) is "selectively smoothed" by a nonlinear diffusion equation, whose diffusion coefficient is inversely related to the degree of rapid increase. This procedure generates a collection of potential pulse time sets ( $\mathbb{T}$ ). Then, via an algorithm which alternates between a Metropolis algorithm on  $\mathbb{T}$  and a time-homogeneous diffusion process on  $\bar{\Theta}$ , a compact manifold with boundary, simulation from an appropriately formulated (posterior) probability measure is achieved. The

---

Received January 27, 2006.

2000 *Mathematics Subject Classification.* Primary 62F15; Secondary 62P10.

*Key words and phrases.* Pulse detection, simulation by diffusion, hormonal secretion, estimation.

Support provided by NSF DMS-0107680 and NIH AG19164, AG19695, AG23133, AG29215, AG14759, DK60717, and M01 RR00585.

*E-mail address:* somesh@stat.fsu.edu

*E-mail address:* dm7b@virginia.edu

*E-mail address:* veldhuis.johannes@mayo.edu

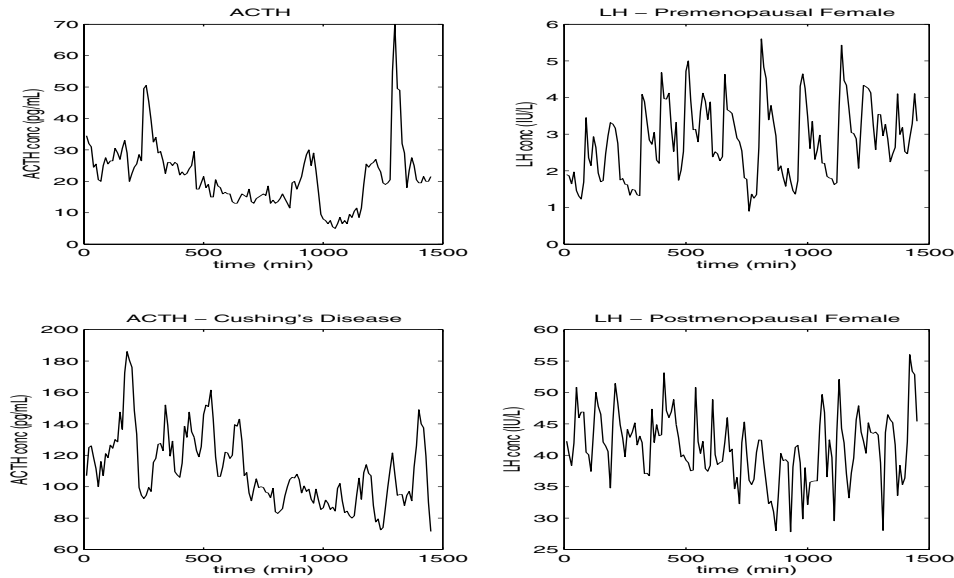


FIG. 1. Hormone concentrations, sampled every 10 min. for 24 hrs. Left: ACTH from a healthy individual and a subject with Cushing's disease. Right: LH from a premenopausal-early phase of menstrual cycle, and a postmenopausal female.

method is applied to recover the underlying structure of brain-pituitary regulation in disease and aging.

**1. Introduction.** Self-adaptive neuroendocrine systems mediate significant control of growth, metabolism, reproduction and stress adaptations. A basic driving property of such ensembles is intermittent signal exchange among brain neurons, the pituitary gland and secreted products of target glands. A fundamental unsolved problem is valid quantification of episodic pituitary output as a window to brain regulatory mechanisms, which cannot be measured directly (Veldhuis [20]). Salient features of pituitary hormone bursts, or pulses, are their (unknown) timing, size and regularity. What are observed are not secretions, but rather assayed concentrations. In Figure 1 are displayed time-sampled pituitary hormone concentrations in the contexts of disease and aging: stress-responsive adrenocorticotropin hormone (ACTH) from both a healthy individual and a subject with Cushing's disease (a pituitary tumor), and the reproductive signal luteinizing hormone (LH) from both pre- and postmenopausal females. The clinician would like to be able to make probability statements about the underlying pulsatile, secretory and kinetic structure of such data. In the present paper we present such a methodology (Table 1, in Section 5).

A probabilistic framework for detecting the locations of the pulse times, without knowing a priori the number or size of the pulses, would require that one do the computations first, conditioned on a given choice, and then vary that choice (i.e., uncondition on the

choice of locations and number). At present, this is computationally infeasible. An alternative is to construct, based upon the data, a collection  $\mathbb{T}$  of decreasing sets, each set being a putative pulse time set. The challenging charge of determining whether and where to add/remove a putative pulse is then directly implementable by moving up or down within this collection. Let  $\bar{\Theta}$  be the parameter space of all the secretory and kinetic parameters of interest. We assume that  $\bar{\Theta} = \Theta \cup \partial\Theta$  is a compact, connected, oriented, finite-dimensional Riemannian manifold with boundary. Manifold structure is assumed for two reasons: first, and most importantly, evidence suggests possible nonlinear dependencies among secretory and kinetic parameters, which future models could then accommodate; secondly, normalized rates of secretion are most naturally described by families of probability densities, and the latter are most appropriately modeled as manifolds. We place a *prior* probability measure on  $S = \mathbb{T} \times \bar{\Theta}$ . Incorporation of the observed data (via a model) then results in a *posterior* probability measure on the parameters of regulated hormone output.

The algorithm consists of two iterated steps. Step 1 (Metropolis): By the Metropolis algorithm, decide to move within  $\mathbb{T}$ , up or down one (add or remove a pulse) or stay the same; Step 2 (Diffusion): “run” (for a fixed amount of time) a time-homogeneous diffusion on  $\bar{\Theta}$ , with reflection in the normal direction at the boundary. Steps 1 and 2 are then iterated, and (asymptotically) the resulting  $\theta_t$  is a realization from the posterior probability measure. For each such realization one can then calculate the value of certain functionals of interest, e.g., fast and slow rates of elimination, total secretion, number of pulses, mass per pulse. Repeating this, one acquires the posterior distribution for these structural entities, and the desired probabilities can then be calculated. Early uses of simulation by diffusion in a linear space setting were presented in Grenander [5], Geman and Hwang [4] and Grenander and Miller [6]. Simulation by diffusion on a compact manifold with boundary was presented by Keenan and Shorter [12].

In the next section we briefly describe the physiology underlying our modeling of pulse (secretory-burst), pulse time, pulse mass, secretion and elimination.

**2. Physiological basis of hormone dynamics.** Physiological stress is a general term used to describe changes that disrupt the body’s homeostasis and elicit defensive immune system responses. At the appropriate time, these defensive responses must be reversed, with homeostasis restored. Cortisol, an adrenal gland steroid, is the body’s primary mechanism for such restoration. Cortisol is controlled by ACTH, which in turn is stimulated by the hypothalamic pulses of corticotropin-release hormone (CRH). Cushing’s disease is the result of a pituitary tumor causing excessive release of ACTH and in turn cortisol, which then greatly weakens the immune state. In male and female reproduction, the steroids testosterone (male) and estrogen and progesterone (female) are transcriptional regulators for a broad range of genes, especially those which are anabolic in nature. They are regulated by the pituitary gland’s luteinizing hormone (LH), with LH in turn being driven by hypothalamic pulses of gonadotropin-releasing hormone (GnRH). Irregular GnRH or LH pulsing results in a loss of bodily anabolic effects in the postmenopausal female and the male, and an inability to ovulate in the

premenopausal female. In the present paper we are concerned with pituitary pulsatile-secreting hormones (e.g., ACTH and LH).

CRH (and the same for GnRH) is secreted by a network of 1000–1500 specialized neurons. Unlike the typical neuron which fires on a 10–100 millisecond scale, the time between firings for these neurons is on the scale of 4–8 minutes. Most of the time, these firings are not synchronized across the network, and the total overall amount of neurotransmitter released is quite small. However, for reasons yet to be explained, on the order of every 45–150 minutes, there is a synchronization of the firing frequencies, at which time significant amounts of the given peptide are secreted (Veldhuis [20]). These occurrences are the previously mentioned hypothalamic pulse (onset) times, with the length of any individual CRH (or GnRH) pulse being quite short, on the order of 1–3 minutes. The small CRH peptide then travels bloodborne to the pituitary gland and stimulates a pulsatile secretion of ACTH (and GnRH stimulates the secretion of LH). CRH and GnRH are initially released into a blood volume of about 1.5 milliliters at the pituitary. When they enter the general circulation (of approximately 5 liters), they are not assayable because of the 3000-fold dilution.

There are multiple mechanisms by which various hormones are degraded and/or removed from the system. In the case of the hypothalamic and pituitary proteins, first, there are enzymes (proteases) in the blood which can reduce various proteins to inactive metabolites. If the degradative cells have appropriate receptors for the given protein, the protein may be internalized by the cell and degraded. A certain amount of protein, depending on the size, charge, degree of glycosylation, may be excreted by the kidneys. Hormones are secreted by their respective glands into interstitial fluid and then into blood, by both advection and diffusion. This occurs quite rapidly (half-life of 1–10 minutes), whereas irreversible degradation is slower (half-life of 15–100 minutes). The two rates of elimination are ordinarily called the fast and slow rates, respectively.

By pulsatile secretion, we mean that release is not a continuous release, but rather at a pulse time, the rate of secretion rapidly increases, followed by a possibly not so rapid decrease. Again, these are general patterns, and one must allow for flexibility, for at times there may be large hypothalamic signals, but little or no pituitary response, or vice-versa. The concepts which need to be quantified are: *pulse (onset) times; pulse mass; normalized rate of pulsatile secretion; basal (non-pulsatile) secretion; and, fast and slow rates of elimination*. These are formulated below.

### 3. Pulsatile-secreted hormone concentration model.

3.1. *Construction of putative pulse time sets:*  $\mathbb{T} = \{\mathcal{T}_N, \dots, \mathcal{T}_p\}$ . In image processing, there is a long history of low-pass filtering being implemented by “running” the heat equation. Perona and Malik [17] improved the edge detection theory by replacing the heat equation by a nonlinear equation of the porous medium type, with data  $u_0$  as the initial datum:

$$\frac{\partial u}{\partial t} = \operatorname{div}(g(|\nabla u|)\nabla u), \quad u(x, 0) = u_0(x), \quad (3.1)$$

where  $g$  is a smooth nonincreasing function with  $g(0) = 1$ ,  $g(x) \geq 0$  and  $g(x)$  tending to 0 at infinity. The condition that  $sg(s)$  is nondecreasing turns out to be a necessary

condition for existence and uniqueness of the solution. Alvarez, Lions and Morel [1] proposed another anisotropic diffusion algorithm for selective smoothing and edge detection. They proposed a partial differential equation of the kind

$$\frac{\partial u}{\partial t} = g(|G * Du|)|Du| \operatorname{div} \frac{Du}{|Du|}, \quad u(0, x, y) = u_0(x, y), \quad (3.2)$$

where  $G$  is a smoothing kernel and  $g$  is a nonincreasing function such that  $\lim_{s \rightarrow \infty} g(s) = 0$ . The term  $|Du| \operatorname{div} \frac{Du}{|Du|} = \Delta u - D^2 u(Du, Du)/|Du|^2$  represents a degenerate diffusion term, which diffuses  $u$  in the direction orthogonal to its gradient  $Du$ . The aim is to smooth along an edge and not across it.

The above is meant to be motivation for our pulse detection method. See Mauger et al. [16] for a review of pulse detection. The two settings, the above and ours, are different in many respects. Fundamental to the above is that the data  $u$  are 2-dimensional, whereas  $g$  is a function of one dimension. In our case, both are 1-dimensional.

For the moment suppose that “data” are given as  $\{Y(t), 0 \leq t \leq 1\}$ , where  $t$  represents the observational time (e.g., over a day). The following algorithm produces a systematic procedure to indicate whether and where to add or remove a pulse time. The equation is run for  $0 \leq s \leq S$ , where  $s$  refers to algorithmic time and  $t$ , again, is observational time:

$$\begin{aligned} u(t, 0) &= Y(t), \quad 0 \leq t \leq 1 && \text{(Initial condition),} \\ u(0, s) &= Y(0), u(1, s) = Y(1), \quad 0 \leq s \leq S && \text{(Dirichlet boundary condition),} \\ \frac{\partial u(t, s)}{\partial s} &= g \left( \left( \frac{\partial u(t, s)}{\partial t} \right)_+ \right) \frac{\partial^2}{\partial t^2} u(t, s) && \text{(Selective smoothing equation),} \\ g(x) &= C_1 / (1 + x/C_2)^2, \quad 0 \leq x \leq C_2, \quad 0 < C_1, \quad 0 < C_2, \end{aligned}$$

where  $(y)_+ = \max(y, 0)$  and  $C_2$  adjusts for the scale. In the present case, the diffusion coefficient  $g(\cdot)$  is a function of the derivative. If the derivative is large positively, there will be very little smoothing at that point ( $t$ ); similarly, if the derivative is positively small or nonnegative, there will be lots of smoothing. The condition that  $xg(x)$  be nondecreasing is satisfied for the above choice. Also, we have tried other boundary conditions (Neumann, mixed), but found the Dirichlet condition to work best. In the discretization of the diffusion (stability conditions given in John [8]), applying it to 10 min. (1/6 hr.) data, we assumed  $\Delta x = (1/6)$ ,  $\Delta t = (1/6)^2$ ,  $C_1 = .003$ ,  $S = 2000$  (usually) and  $C_2$  is the maximum positive derivative (difference) calculated for the data.

Heuristically, the algorithm is as follows. Suppose that the time series  $Y$  has  $N$  local minima. As the selective smoothing proceeds, one of the local minima (‘the weakest’) will be smoothed away and the set of local minima will comprise  $N - 1$  points. As the smoothing continues, additional local minima will be removed. In practice, some pulses evolve with a stuttering onset, wherein an initial slight increase precedes a large rapid increase; in the present method such points are not excluded from putative pulse-time sets. For pragmatic implementation, smoothing evolves for some pre-specified number of algorithmic cycles or until some pre-specified minimal number (e.g.,  $p$ ) of pulse times is attained. The result is the collection of decreasing pulse time sets:  $\mathbb{T} = \{\mathcal{T}_N, \dots, \mathcal{T}_p\}$ . In Figure 2 the constructed surfaces ( $u$ ) are displayed, and in Figure 3 the four  $\mathbb{T}$  collections

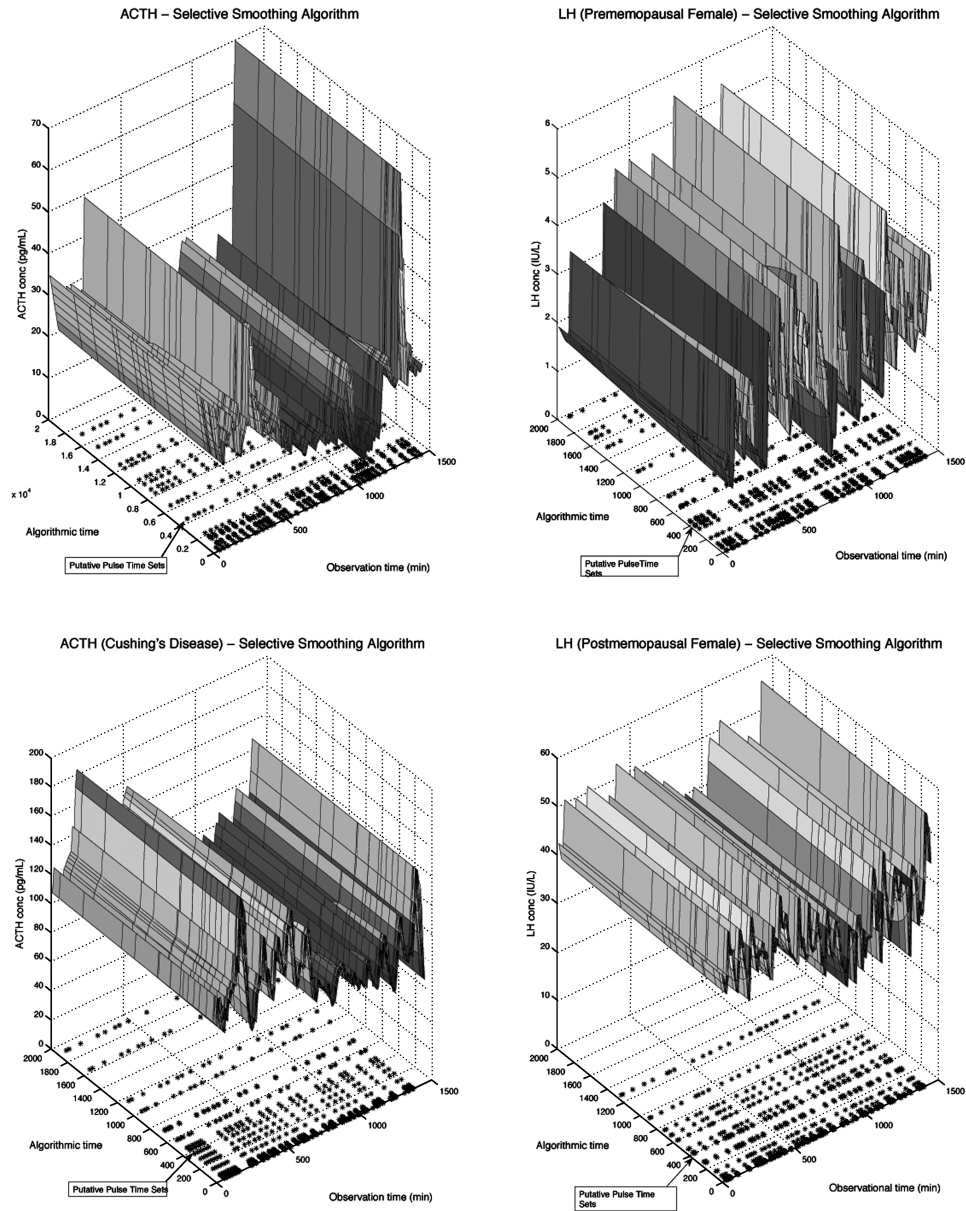


FIG. 2. Application of Selective Smoothing Algorithm, determining Putative Pulse Time Sets. The four subpanels correspond to the four profiles of Figure 1. The asterisks (\*) denote the pulse times. The pulse time sets are progressively decreasing as algorithmic time proceeds.

are displayed so that one can easily follow the progression from the maximum set to the minimum set.

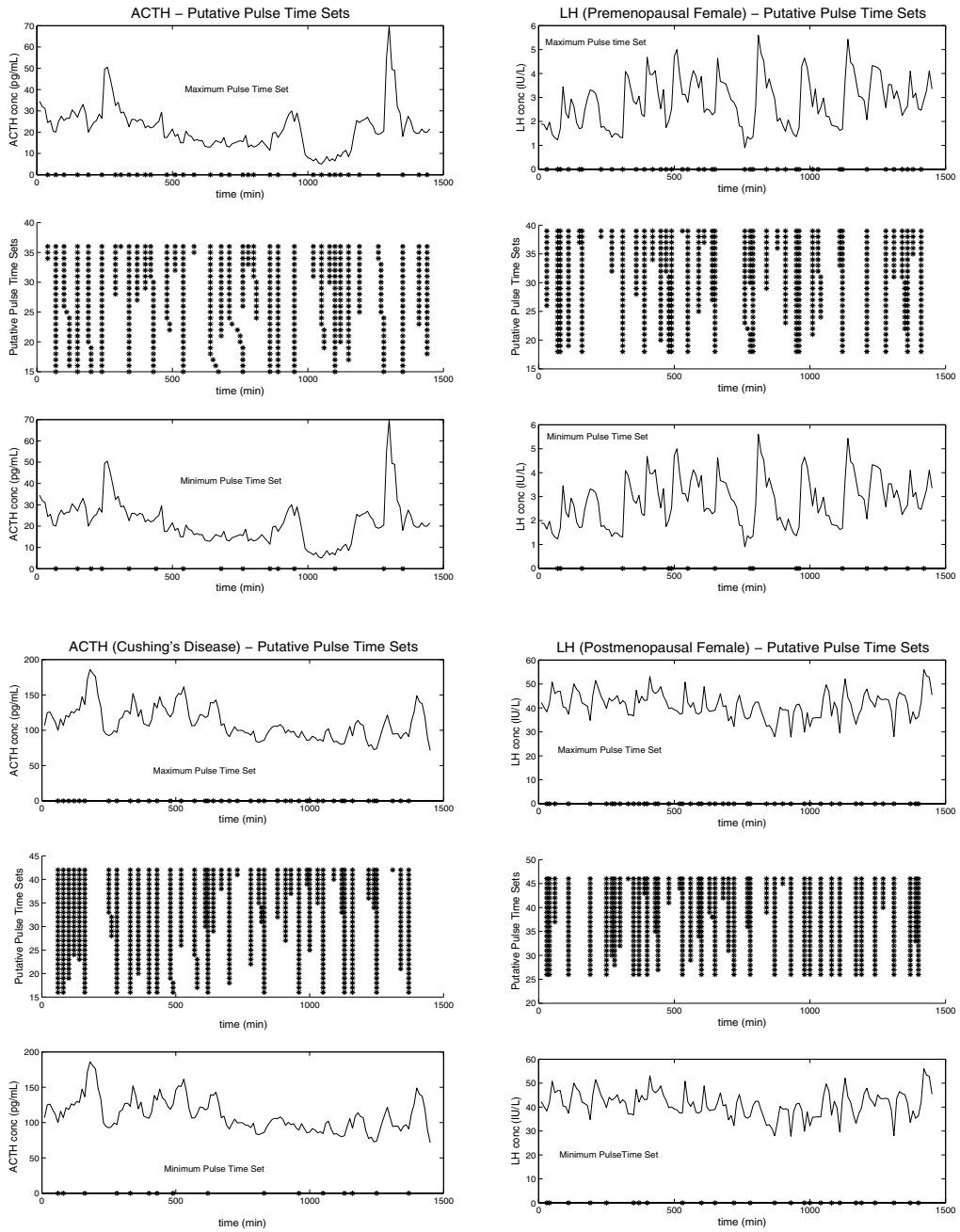


FIG. 3. The Putative Pulse Time Sets, from Figure 2, displayed in a more revealing manner. There are three subpanels for each: the top and bottom subpanels show the relationship of the maximum and minimum pulse sets to the concentration profile; the middle subpanel shows the progression (as pulse times are removed) from the maximum set to the minimum set.

3.2. *A model of concentration for a given pulse times set  $\mathcal{T}_m$ .* In Keenan, Veldhuis, and Yang [15], Keenan and Veldhuis [14], and Keenan, Sun and Veldhuis [13], a model of hormone secretion and concentration is presented, conditioned on the set of pulse times  $\mathcal{T}_m = (T_1, \dots, T_m)$ . A biexponential elimination rate is used, where the fast rate  $\alpha_1$  represents advection and diffusion, and the slower rate  $\alpha_2$  represents elimination. The fractional weights between the fast and slow rates,  $a^{(1)}$  and  $a^{(2)} = 1 - a^{(1)}$  have been reasonably well determined by clinical experiments.

Each pulse time  $T_j$  marks the onset of a secretory burst. Conceptually, a burst reflects abrupt exocytotic discharge of hormone-containing granules followed by less rapid secretion of newly synthesized molecules. The mass of hormone secreted in any given burst is assumed to be the sum of a finite amount of minimally available stores ( $\eta_0$ ), a linear function ( $\eta_1$ ) of hormone accumulation over the preceding interpulse interval, and an allowable flexibility (a random element  $A_j$ ) in individual burst mass,

$$M_j = \eta_0 + \eta_1 \times (T_j - T_{j-1}) + A_j. \quad (3.3)$$

The  $A_j$ 's are I.I.D.  $N(0, \sigma_A^2)$  and describe the variation in pulse mass due to nonuniform cellular release. The mass contained in any given burst,  $M_j$ , is released according to an adaptable (hormone-, subject- and condition-specific) waveform. The waveform (evolution of the instantaneous secretion rate over time) is represented via the three-parameter generalized Gamma (probability) density, which encapsulates the normalized rate of secretion (mass units) over time (minutes) per unit distribution volume (L):

$$\psi(r) = \frac{\beta_3}{\Gamma(\beta_1)\beta_2^{\beta_1\beta_3}} r^{(\beta_1\beta_3)-1} e^{-(r/\beta_2)^{\beta_3}}. \quad (3.4)$$

This pulsatile release is superimposed on a continuous basal release  $\beta_0$ . The instantaneous secretion rate  $Z(r)$  and the resulting concentrations  $X(t)$  are then given as:

$$Z(r) = \beta_0 + \sum_{T_j \leq r} M_j \psi(r - T_j),$$

$$X(t) = (a^{(1)}e^{-\alpha_1 t} + a^{(2)}e^{-\alpha_2 t})X(0) + \int_0^t (a^{(1)}e^{-\alpha_1(t-s)} + a^{(2)}e^{-\alpha_2(t-s)})Z(s) ds.$$

The secretory and kinetic parameter is  $\theta = (\gamma, \sigma) = ((\beta_0, \alpha_1, \alpha_2, \eta_0, \eta_1, \beta_1, \beta_2, \beta_3), (\sigma_A, \sigma_\epsilon))$ , and  $\bar{\Theta}$  represents the resulting parameter space. Estimation of the various parameters is then implemented by way of a discretization of the model; the details are given in Keenan, Veldhuis, and Yang [15], Keenan, Licinio and Veldhuis [11], Yang [21], and Keenan, Alexander et al [9]. The discretization results in  $Y = (Y_1, Y_2, \dots, Y_n)'$  having a normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$ , which are functions of  $\theta$  and the pulse times  $\mathcal{T}_m$ . The log-likelihood and the (minus) log-posterior density then have the following forms:

$$l_{\mathcal{T}_m}(\theta) = \log L(\theta) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \log |\Sigma| - \frac{1}{2} (Y - \mu)' \Sigma^{-1} (Y - \mu),$$

$$H(\mathcal{T}_m, \theta) \propto -l_{\mathcal{T}_m}(\theta) - \log(\pi_1(\mathcal{T}_m) \times \pi_2(\theta)), \quad (\mathcal{T}_m, \theta) \in \mathbb{T} \times \bar{\Theta},$$

where  $\pi_1$  is the prior probability density on  $\mathbb{T}$ ,  $\pi_2$  is the prior probability density on  $\bar{\Theta}$  and  $e^{-H}$  is the posterior density. In applications we use a uniform (constant) prior on



$\bar{\Theta}$ , and the prior on  $\mathbb{T}$  is  $\pi_2(\mathcal{T}_m) \propto e^{-m}$ , which corresponds to an Akaike Information Criterion (AIC) penalization on the number of pulses.

**4. Justification of the algorithm.** We are given a function  $H(\cdot, \cdot)$  defined over  $\mathbb{T} \times \bar{\Theta}$ , where  $\mathbb{T}$  is a discrete set  $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_N\}$  and  $\bar{\Theta}$  is a compact, regular domain (see Boothby [2]), in the sense that there exists a  $q$ -dimensional orientable manifold  $Q$  of class  $C^\infty$ , with  $\Theta$  being a  $d$ -dimensional domain in  $Q$  and whose closure  $\bar{\Theta}$  is compact and whose boundary  $\partial\Theta = \bar{\Theta} - \Theta$  consists of a finite number of hypersurfaces each of dimension  $d - 1$  and of class  $C^3$ . Let  $\bar{\Theta}$  be a compact, connected, oriented, smooth  $d$ -dimensional Riemannian manifold with boundary and with metric  $g$ . We assume an ordering  $\prec$  in  $\mathbb{T}$ , namely,  $\mathcal{T}_1 \prec \mathcal{T}_2 \prec \dots \prec \mathcal{T}_N$ . In the algorithm the order is the subset order. We also assume that  $H(\mathcal{T}_i, \cdot)$  is a twice continuously differentiable function for each fixed  $\mathcal{T}_i$ ,  $1 \leq i \leq N$ . Let  $C^\infty(\bar{\Theta})$  denote the class of real-valued infinitely differentiable functions on  $\bar{\Theta}$ . Let  $\Omega$  be the natural volume element of the oriented Riemannian manifold  $\bar{\Theta}$  and let  $\lambda$  be the Borel measure associated with  $\Omega$ . Integration of functions on  $\bar{\Theta}$  (which is integration of  $d$ -forms on  $\bar{\Theta}$ ) can be viewed as Lebesgue-Stieltjes integration on  $\bar{\Theta}$  with respect to  $\lambda$ :

$$\int_{\bar{\Theta}} f = \int_{\bar{\Theta}} f \Omega = \int_{\bar{\Theta}} f(\theta) \lambda(d\theta) \quad \text{for all } f \in C(\bar{\Theta}).$$

Let  $\nabla$  and  $\Delta$  denote the gradient and Laplacian on  $\bar{\Theta}$ , with respect to  $g$ . For a broad class of operators  $A$ , including the particular case of  $A = \frac{1}{2}(\Delta - \nabla h) = \frac{1}{2}\Delta - \frac{1}{2}\nabla h$ , where  $h \in C^\infty(\bar{\Theta})$ , S. Ito [7] considered the behavior of the fundamental solution to the Cauchy problem for the following parabolic equation. For a given  $f \in C(\bar{\Theta})$ , find  $v$  such that

$$\begin{aligned} \frac{\partial}{\partial t} v(t, x) &= A v(t, x), \quad t > 0, x \in \bar{\Theta}, \\ Lv(t, x) &= \frac{\partial}{\partial n} v(t, x) = 0, \quad t > 0, x \in \partial\bar{\Theta} \quad (\text{Boundary condition (reflecting barrier)}), \\ \lim_{t \downarrow 0} v(t, x) &= f(x) \quad (\text{uniformly in } x \in \bar{\Theta}) \quad (\text{Initial condition}). \end{aligned}$$

It was shown that the fundamental solution is jointly continuous, and under the above boundary condition (reflection at the boundary in the normal direction), is strictly positive:  $p(t, x, y) > 0$ ,  $x, y \in \bar{\Theta}$ ,  $t > 0$ . In Sato and Ueno [19], it is shown that this fundamental solution is the transition density for the time-homogeneous Markov (Feller) process with infinitesimal generator:  $A = \frac{1}{2}\Delta - \frac{1}{2}\nabla h$ , where there is reflection at the boundary. (In a linear space, the diffusion is  $dX_t = -\frac{1}{2}\nabla h(X_t)dt + dW_t$ .) If the reflection were not in the normal direction, the Markov process would spend local time on the boundary, and the density would not necessarily be strictly positive. In Keenan and Shorter [12], in the context of a compact manifold with boundary, it is shown that the process  $\{X_t, t \geq 0\}$  converges weakly to the measure  $e^{-h}$ . For the present setting,  $h(\cdot) = H(\mathcal{T}_k, \cdot)$  plus a constant (so that  $e^{-h}$  is the posterior conditional density of  $\theta$  given  $\mathcal{T}_k$ ). What is most important for our concern is that the transition density is strictly positive:  $p(t, x, y) > 0$ ,  $x, y \in \bar{\Theta}$ ,  $t > 0$ . Let  $S$  denote the product space  $\mathbb{T} \times \bar{\Theta}$ , and  $\mathcal{F}$  and  $\mathcal{S}$  the Borel  $\sigma$ -fields of subsets on  $\bar{\Theta}$  and  $S$ . Let  $\mathcal{P}$  denote the powerset of  $\mathbb{T}$ . By

$\lambda$ ,  $c$  and  $\nu$  we denote, respectively, the Borel measure associated with the natural volume element on  $(\overline{\Theta}, \mathcal{F})$ , counting measure on  $(\mathbb{T}, \mathcal{P})$  and the product measure  $c \times \lambda$  on  $(S, \mathcal{S})$ .

We have two steps in each iteration of the algorithm. The first step is a jump in the first argument of  $H(\cdot, \cdot)$  where we keep the second argument fixed. The second step is a diffusion in the second argument of  $H(\cdot, \cdot)$  where we keep the first argument fixed. We denote by  $(\mathcal{T}^{(k)}, \theta^{(k)})$  the point in  $\mathbb{T} \times \overline{\Theta}$  after  $k$  iterations. The starting point is denoted by  $(\mathcal{T}^{(0)}, \theta^{(0)})$ . In the  $k$ -th iteration (for  $k \geq 1$ ) we have the following two steps.

**Step 1. (Jump):** Suppose  $\mathcal{T}^{(k-1)} = \mathcal{T}_i$ . If  $2 \leq i \leq N-1$ , then, as in the Metropolis algorithm, either we move to  $\mathcal{T}_{i-1}$  with probability  $\xi \min\{1, \exp[-(H(\mathcal{T}_{i-1}, \theta^{(k-1)}) - H(\mathcal{T}_i, \theta^{(k-1)}))]\}$  or to  $\mathcal{T}_{i+1}$  with probability  $\xi \min\{1, \exp[-(H(\mathcal{T}_{i+1}, \theta^{(k-1)}) - H(\mathcal{T}_i, \theta^{(k-1)}))]\}$  or we stay at  $\mathcal{T}_i$ . Here  $\xi$  is a constant with  $0 < \xi \leq \frac{1}{2}$ . We denote by  $\mathcal{T}^{(k)}$  the new  $\mathcal{T}$  chosen in this way after the first step of the  $k$ -th iteration. So, we have the following transition probabilities for a fixed value  $\theta$  of  $\theta^{(k-1)}$ . For  $2 \leq i \leq N-1$ ,

$$\begin{aligned} P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_{i-1} \mid \mathcal{T}^{(k-1)} = \mathcal{T}_i) &= \xi \min\{1, \exp[-(H(\mathcal{T}_{i-1}, \theta) - H(\mathcal{T}_i, \theta))]\}, \\ P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_i \mid \mathcal{T}^{(k-1)} = \mathcal{T}_i) &= 1 - \xi \min\{1, \exp[-(H(\mathcal{T}_{i-1}, \theta) - H(\mathcal{T}_i, \theta))]\} \\ &\quad - \xi \min\{1, \exp[-(H(\mathcal{T}_{i+1}, \theta) - H(\mathcal{T}_i, \theta))]\}, \\ P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_{i+1} \mid \mathcal{T}^{(k-1)} = \mathcal{T}_i) &= \xi \min\{1, \exp[-(H(\mathcal{T}_{i+1}, \theta) - H(\mathcal{T}_i, \theta))]\}. \end{aligned}$$

We need to modify the above algorithm if  $i$  is 1 or  $N$ . If  $i = 1$ , we imagine  $\mathcal{T}_0$  as identically equal to  $\mathcal{T}_1$ . Then we have the following transition probabilities:

$$\begin{aligned} P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_1 \mid \mathcal{T}^{(k-1)} = \mathcal{T}_1) &= 1 - \xi \min\{1, \exp[-(H(\mathcal{T}_2, \theta) - H(\mathcal{T}_1, \theta))]\}, \\ P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_2 \mid \mathcal{T}^{(k-1)} = \mathcal{T}_1) &= \xi \min\{1, \exp[-(H(\mathcal{T}_2, \theta) - H(\mathcal{T}_1, \theta))]\}. \end{aligned}$$

When  $i = N$ , we modify the transition probabilities in a similar way. For  $|i - j| > 1$ ,

$$P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_j \mid \mathcal{T}^{(k-1)} = \mathcal{T}_i) = 0.$$

**Step 2. (Diffusion):** For the fixed  $\mathcal{T}^{(k)}$  obtained in the above procedure, we run a diffusion for  $(k-1)t_0 \leq t \leq kt_0$  with reflecting boundary. The value of  $t_0$  is fixed and independent of  $k$ . The diffusion is that described above, with infinitesimal generator  $A = \frac{1}{2}\Delta - \frac{1}{2}\nabla h$ ,  $h(\cdot) = H(\mathcal{T}_k, \cdot)$  plus a constant, and there is reflection in the inward normal direction at the boundary. By  $\theta^{(k)}$  we denote  $\theta_{kt_0}$  obtained in this way. This completes one iteration.

One then repeats the above two steps alternately. By  $X_k$  we will denote the pair  $(\mathcal{T}^{(k)}, \theta^{(k)})$  obtained using the above algorithm. Then  $\{X_k : k \geq 0\}$  is a discrete time parameter stochastic process. This process is a time homogeneous Markov chain with state space  $\mathbb{T} \times \overline{\Theta}$ . The Markov property of  $X_k$  is easy to prove since we are defining the process by transition from  $X_{k-1}$  to  $X_k$ , which is not dependent on the values of  $X$  before time  $k-1$  except through the value of  $X_{k-1}$ . Time homogeneity is also clear from the fact that the transition probabilities from  $X_{k-1}$  to  $X_k$  do not depend on the value of  $k$ . Being a time homogeneous Markov process, the diffusion in Step 2 has a transition probability function  $Q(\mathcal{T}^{(k)}, t, \theta, B)$ , which is the probability for going from  $\theta \in \overline{\Theta}$  to  $B \subset \overline{\Theta}$  if we run the diffusion for time  $t$ .

From the results of Sato and Ueno [19] and Ito [7] that we have discussed earlier, it follows that: For each fixed  $\mathcal{T}^{(k)}$  and  $t$ ,  $Q(\mathcal{T}^{(k)}, t, \theta, \cdot) \ll \lambda$  and  $q(\mathcal{T}^{(k)}, t, \theta, \theta')$  is positive for all  $\theta, \theta' \in \overline{\Theta}$  and continuous in  $(\theta, \theta')$ , where  $q(\mathcal{T}^{(k)}, t, \theta, \cdot)$  denotes the density of  $Q(\mathcal{T}^{(k)}, t, \theta, \cdot)$  with respect to  $\lambda$ .

Let  $P_n(x, A)$  denote the  $n$ -step transition probability of the process  $\{X_k : k \geq 0\}$ , where  $x \in S$  and  $A \in \mathcal{S}$ . When  $n = 1$ , we write the transition probability simply as  $P(x, A)$ . Any  $\mathcal{S}$ -measurable set  $A$  can be written as a union of disjoint sets  $A_1, A_2, \dots, A_N$ ,  $A = \bigcup_{i=1}^N A_i$ , where  $A_i = A \cap (\{\mathcal{T}_i\} \times \overline{\Theta}) = \{\mathcal{T}_i\} \times B_i$  for some  $B_i \in \mathcal{F}$  for  $1 \leq i \leq N$ . Therefore, if  $\nu(A) = 0$ , then  $\lambda(B_i) = 0$  for  $1 \leq i \leq N$  and hence  $Q(\mathcal{T}_i, t_0, \theta', B_i) = 0$  for  $1 \leq i \leq N$  since  $Q(\mathcal{T}_i, t_0, \theta', \cdot) \ll \lambda$  for  $1 \leq i \leq N$ . Hence,  $\nu(A) = 0$  implies  $P(x, A) = 0$  for each  $x \in S$ , and  $P(x, \cdot) \ll \nu$  for all  $x \in S$ . By induction, since  $P_n(x, A) = \int_S P_{n-1}(x, dy)P(y, A) = 0$  for all  $x \in S$  and  $n \geq 2$ , it follows that  $P_n(x, \cdot) \ll \nu$  for each  $x \in S$  and  $n \geq 1$ . Let  $p_n(x, y)$  denote the density of  $P_n(x, \cdot)$  with respect to  $\nu$ . When  $n = 1$ , we write the density simply as  $p(x, y)$ . Thus, we have the transition density of the process  $\{X_k : k \in \mathbb{N}\}$  as  $p((\mathcal{T}_i, \theta), (\mathcal{T}_j, \theta')) = P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_j \mid \mathcal{T}^{(k-1)} = \mathcal{T}_i) \cdot q(\mathcal{T}_j, t_0, \theta, \theta')$  and for  $n > 1$ , we have the  $n$ -step transition density recursively defined as

$$p_n((\mathcal{T}_i, \theta), (\mathcal{T}_j, \theta'')) = \sum_{l=1}^N \int_{\overline{\Theta}} p_{n-1}((\mathcal{T}_i, \theta), (\mathcal{T}_l, \theta')) \cdot p((\mathcal{T}_l, \theta'), (\mathcal{T}_j, \theta'')) \lambda(d\theta'). \quad (4.1)$$

The following lemma will be useful in the course of proving  $\nu$ -irreducibility and aperiodicity of the Markov chain  $X_k$ .

LEMMA 4.1. Let  $A$  be a compact set in  $\overline{\Theta}$ . Suppose  $f$  and  $g$  are two continuous functions from  $A \times A$  to  $R$ . Define a function  $h : A \times A \rightarrow R$  as  $h(x, y) = \int_A f(x, z)g(z, y)\lambda(dz)$ . Then  $h$  is uniformly continuous on  $A \times A$ .

*Proof.* Straightforward, elementary analysis argument. □

LEMMA 4.2. For fixed integers  $i, j$  with  $1 \leq i \leq N$  and  $1 \leq j \leq N$ , and a positive integer  $n$ , (i)  $p_n((\mathcal{T}_i, \theta), (\mathcal{T}_j, \theta'))$  is positive for all  $\theta, \theta' \in \overline{\Theta}$  and continuous in  $(\theta, \theta')$  if  $|i - j| = n$ , (ii)  $p_n((\mathcal{T}_i, \theta), (\mathcal{T}_j, \theta')) = 0$  for all  $\theta, \theta' \in \overline{\Theta}$  if  $|i - j| > n$ , and (iii) for  $1 \leq i \leq N$ ,  $\inf_{\theta, \theta' \in \overline{\Theta}} p_2((\mathcal{T}_i, \theta), (\mathcal{T}_i, \theta')) > 0$ .

*Proof.* Results (i) and (ii) hold for  $n = 1$ , since  $P_\theta(\mathcal{T}^{(k)} = \mathcal{T}_j \mid \mathcal{T}^{(k-1)} = \mathcal{T}_i)$  is positive for  $|i - j| = 1$  and zero for  $|i - j| > 1$  and for each  $\mathcal{T}^{(k)}$  and  $t$ ,  $q(\mathcal{T}^{(k)}, t, \theta, \theta')$  is continuous in  $(\theta, \theta')$  and positive for all  $\theta, \theta' \in \overline{\Theta}$ . The proof is by induction on  $n$ . Result (iii) is a consequence of compactness and Lemma 4.1. □

DEFINITION 4.3. Suppose  $\{X_n : n \geq 0\}$  is a time homogeneous Markov chain with the state space  $S$  and  $\mathcal{S}$  is a  $\sigma$ -field of subsets of  $S$  and  $P_n(x, A)$  denotes  $n$ -step transition probability from  $x \in S$  to  $A \in \mathcal{S}$ .  $\{X_n : n \geq 0\}$  is called irreducible with respect to a measure  $\mu$  on  $\mathcal{S}$  if for each  $A \in \mathcal{S}$  with  $\mu(A) > 0$  and for each  $x \in S$ , there exists  $n \geq 1$  such that  $P_n(x, A) > 0$ .

THEOREM 4.4. The process  $X_n$  is irreducible with respect to the measure  $\nu$ .

*Proof.* Any  $\mathcal{S}$ -measurable set  $A$  can be written as  $A = \bigcup_{i=1}^N A_i$ , a union of disjoint sets  $A_1, A_2, \dots, A_N$  where  $A_i = A \cap (\{\mathcal{T}_i\} \times \overline{\Theta})$  for  $1 \leq i \leq N$ . Therefore, each  $A_i$  can be written as  $A_i = \{\mathcal{T}_i\} \times B_i$  for some  $B_i \in \mathcal{F}$ . Since  $\nu(A) = \sum_{i=1}^N \lambda(B_i)$ ,  $\nu(A) > 0$  implies  $\sum_{i=1}^N \lambda(B_i) > 0$ , which in turn implies that at least one of  $\lambda(B_i)$ 's must be positive; assume that  $\lambda(B_j) > 0$ . Suppose  $x = (\mathcal{T}_l, \theta)$  and  $m = |l - j|$ . If  $m > 0$ , it follows from the first part of Lemma 4.2 that  $P_m(x, A) > 0$ . If  $m = 0$ , from the third part of Lemma 4.2 it follows that  $P_2(x, A) > 0$ .  $\square$

To prove Harris recurrence, we introduce some notation and results, found in Revuz [18]. Let  $P$  denote the operator associated with the transition probability defined as  $P(f)(x) = \int_S P(x, dy)f(y)$ , for measurable functions  $f$  on  $(S, \mathcal{S})$  for which at least one of  $P(f^+)$  and  $P(f^-)$  is finite. For  $h : (S, \mathcal{S}) \rightarrow [0, 1]$ , we define the operator  $U_h$  as  $U_h = \sum_{n=0}^{\infty} (PM_{1-h})^n P = \sum_{n=0}^{\infty} P(M_{1-h}P)^n$ , where  $M_h$  denotes the operator of multiplication by the function  $h$ . If  $h$  is a constant functions, i.e.,  $h \equiv c \in [0, 1]$ , we denote  $U_h$  by  $U_c$ . One can associate the operator  $U_h$  with a kernel  $U_h(\cdot, \cdot)$ , which is defined as  $U_h(x, A) = U_h 1_A(x)$  for  $x \in S$  and  $A \in \mathcal{S}$ . Then  $U_h f(x) = \int_S U_h(x, dy)f(y)$ . We will use the following resolvent equation in the course of proving Harris recurrence of the chain  $X_k$ . For any two  $\mathcal{S}$ -measurable functions  $h_1$  and  $h_2$  from  $S$  to  $[0, 1]$  with  $h_1 \leq h_2$ ,  $U_{h_1} = U_{h_2} + U_{h_1}M_{h_2-h_1}U_{h_2} = U_{h_2} + U_{h_2}M_{h_2-h_1}U_{h_1}$ .

LEMMA 4.5. (i) For  $0 < c < 1$ ,  $U_c(x, A) = \sum_{n \geq 1} (1 - c)^{n-1} P_n(x, A)$  for all  $x \in S$  and  $A \in \mathcal{S}$ ; (ii) for each  $c \in (0, 1)$  and  $x \in S$ ,  $U_c(x, \cdot) \ll \nu$  and the density  $b_c(x, y)$  of  $U_c(x, \cdot)$  with respect to  $\nu$  is given by  $b_c(x, y) = \sum_{n \geq 1} (1 - c)^{n-1} p_n(x, y)$ ; and (iii) for all  $c \in (0, 1)$ ,  $\beta_c = \inf_{x, y \in S} b_c(x, y)$  is strictly positive.

*Proof.* Straightforward.  $\square$

LEMMA 4.6. There is a constant  $h_0 \in (0, 1)$  and a positive measure  $m_0$  equivalent to  $\nu$  such that  $U_{h_0} \geq U_{h_0}(h_0) \otimes m_0$ , where  $\otimes$  is defined as  $(a \otimes \mu)f(x) = a(x) \int_E f(y)\mu(dy)$ , for a measurable function  $a$  on  $E$  and a measure  $\mu$  on a  $\sigma$ -field  $\mathcal{E}$  of subsets of  $E$ .

*Proof.* Let  $0 < c' < c'' < 1$ . By part (iii) of Lemma 4.5,  $U_{c'}U_{c''} \geq \beta_{c'}\beta_{c''}\nu(S) \otimes \nu$ . Set  $h_0 = \min(\frac{1}{2}c', \beta_{c'}\beta_{c''})$ . Then by the resolvent equation,

$$\begin{aligned} U_{c'} &= U_{c''} + (c'' - c')U_{c'}U_{c''} \geq (c'' - c')U_{c'}U_{c''} \geq (c'' - c')\beta_{c'}\beta_{c''}\nu(S) \otimes \nu \\ &\geq (c'' - c')\nu(S) \cdot h_0 \otimes \nu. \end{aligned}$$

Again by the resolvent equation,

$$U_{h_0} \geq U_{h_0}M_{c'-h_0}U_{c'} \geq \frac{1}{2}c'U_{h_0}U_{c'} \geq \frac{1}{2}c'(c'' - c')\nu(S)U_{h_0}(h_0) \otimes \nu.$$

Now we set  $m_0 = \frac{1}{2}c'(c'' - c')\nu(S)\nu$  and the result of the lemma follows.  $\square$

RESULT 4.7 (Revuz [18]). A Markov chain  $\{Y_n\}$  on  $(S, \mathcal{S})$  is recurrent in the sense of Harris if there is a strictly positive  $\mathcal{S}$ -measurable function  $h_0$  and a nonzero measure  $m_0$ , such that  $U_{h_0}(h_0) = 1$  on  $S$  and  $U_{h_0} \geq 1 \otimes m_0$ .

THEOREM 4.8. The Markov chain  $\{X_k : k \geq 0\}$  is recurrent in the sense of Harris.

*Proof.* By Lemma 4.6, there is a constant  $h_0 \in (0, 1)$  and a positive measure  $m_0$  equivalent to  $\nu$  such that  $U_{h_0} \geq U_{h_0} h_0 \otimes m_0$ . Since  $h_0$  is a constant and  $h_0 \in (0, 1)$ , we have for  $x \in S$ ,

$$\begin{aligned} U_{h_0} h_0 &= E_x \left[ \sum_{n \geq 1} (1 - h_0)^{n-1} h_0 \right] = \sum_{n \geq 1} \{ (1 - h_0)^{n-1} - (1 - h_0)^n \} \\ &= 1 - \lim_{n \rightarrow \infty} (1 - h_0)^n = 1. \end{aligned}$$

Hence,  $U_{h_0} \geq 1 \otimes m_0$ . Therefore, by Result 4.7,  $X_k$  is Harris recurrent.  $\square$

RESULT 4.9 (Revuz [18]). For a Harris chain with  $n$ -step transition probability  $P_n(\cdot, \cdot)$  there are only two possibilities:

(i) either for any pair  $(\nu_1, \nu_2)$  of probability measures on  $S$ ,

$$\lim_{n \rightarrow \infty} \|\nu_1 P_n - \nu_2 P_n\| = 0, \tag{4.2}$$

(ii) or there exists an integer  $d \geq 2$  and  $\mathcal{S}$ -measurable sets  $C_1, C_2, \dots, C_d, F$  of  $S$  such that  $\bigcup_{i=1}^d C_i \cup F = S$ ,  $\nu(F) = 0$  and  $P(x, C_{i+1}) = 1$  for all  $x \in C_i$  for  $i = 1, \dots, d - 1$  and  $P(x, C_1) = 1$  for all  $x \in C_d$ .

DEFINITION 4.10. In the first case of Result 4.9 the chain is called aperiodic. In the second case the chain is called periodic and the integer  $d$  is called the period.

THEOREM 4.11. The Markov chain  $X_k$  is aperiodic.

*Proof.* By Theorem 4.8,  $X_k$  is Harris recurrent. We will show that the second case in Result 4.9 cannot occur. We will prove it by contradiction.

Denote  $\{\mathcal{T}_1\} \times \overline{\Theta}$  by  $S_1$ . Suppose the second case of Result 4.9 holds. Then  $S_1 \cap (\bigcup_{j=1}^d C_j)$  is not empty because otherwise  $S_1 \subset F$ , which would imply that  $\nu(S_1) = 0$  contradicting the fact that  $\nu(S_1) = \lambda(\overline{\Theta}) > 0$ . Suppose  $x \in S_1 \cap (\bigcup_{j=1}^d C_j)$ . Then  $x \in C_j$  for some  $j$ . Without loss of generality assume that  $x \in C_d$ . Then  $P(x, C_1) = 1$  by the assumption of the second case of Result 4.9. Therefore,  $P(x, C_1^c) = 0$  and hence  $P(x, S_1 \cap C_1^c) = 0$ . We assert that  $\nu(S_1 \cap C_1^c) = 0$ . We prove this assertion by contradiction. Suppose  $\nu(S_1 \cap C_1^c) > 0$ . Then  $S_1 \cap C_1^c = \{\mathcal{T}_1\} \times B$  for some  $B \in \mathcal{F}$  and  $\lambda(B) = \nu(S_1 \cap C_1^c) > 0$ . Now  $p((\mathcal{T}_1, \theta), (\mathcal{T}_1, \theta')) > 0$  for all  $\theta, \theta' \in \overline{\Theta}$  because  $P_\theta(\mathcal{T}^{(1)} = \mathcal{T}_1 \mid \mathcal{T}^{(0)} = \mathcal{T}_1) > 0$  for all  $\theta \in \overline{\Theta}$  and  $q(\mathcal{T}_1, t_0, \theta, \theta') > 0$  for all  $\theta, \theta' \in \overline{\Theta}$ . Therefore,

$$P(x, S_1 \cap C_1^c) = \int_{S_1 \cap C_1^c} p(x, (\mathcal{T}_1, \theta')) \lambda(d\theta') > 0.$$

Therefore, our assertion that  $\nu(S_1 \cap C_1^c) = 0$  is proved. Hence  $\nu(S_1 \cap C_1) = \nu(S_1) > 0$ . Therefore,  $S_1 \cap C_1$  is nonempty. So, there exists  $y \in S_1 \cap C_1$ . Again, by the assumption of the second case of Result 4.9,  $P(y, C_2) = 1$ . Similar arguments show that  $\nu(S_1 \cap C_2^c) = 0$ . So we have

$$\nu(S_1 \cap (C_1 \cap C_2)^c) \leq \nu(S_1 \cap C_1^c) + \nu(S_1 \cap C_2^c) = 0.$$

So,  $\nu(S_1 \cap (C_1 \cap C_2)^c) = 0$  and hence  $\nu(S_1 \cap (C_1 \cap C_2)) = \nu(S_1) > 0$ . This is a contradiction to our assumption that  $C_1$  and  $C_2$  are disjoint since in that case we would have  $\nu(S_1 \cap (C_1 \cap C_2)) = 0$ . This completes the proof.  $\square$

THEOREM 4.12. The Markov chain  $X_k$  has an invariant probability measure  $\pi$  which is absolutely continuous with respect to the measure  $\nu = \lambda \times c$ . The density of  $\pi$  with respect to  $\nu$  is given by

$$g(\mathcal{T}, \theta) = \frac{\exp(-H(\mathcal{T}, \theta))}{\sum_{i=1}^N \int_{\overline{\Theta}} \exp(-H(\mathcal{T}_i, \theta')) \lambda(d\theta')}. \quad (4.3)$$

*Proof.* Any set  $A \in \mathcal{S}$  can be written as a disjoint union  $A = \bigcup_{i=1}^N A_i$  of sets, where  $A_i = A \cap (\{\mathcal{T}_i\} \times \overline{\Theta})$  for  $1 \leq i \leq N$ . Then  $A_i = \{\mathcal{T}_i\} \times B_i$  for some  $B_i \in \mathcal{F}$ .

So, the transition probability  $P((\mathcal{T}_i, \theta), A_j) = 0$  if  $|i - j| > 1$ . For  $(\mathcal{T}_i, \theta) \in S$ ,  $2 \leq i \leq N - 1$  and  $A \in \mathcal{S}$ , the transition probability of  $X_k$  can be written as

$$\begin{aligned} P((\mathcal{T}_i, \theta), A) &= \xi \min\{1, \exp[-(H(\mathcal{T}_{i-1}, \theta) - H(\mathcal{T}_i, \theta))]\} \int_{B_{i-1}} q(\mathcal{T}_{i-1}, t_0, \theta, \theta') \lambda(d\theta') \\ &\quad + \xi \min\{1, \exp[-(H(\mathcal{T}_{i+1}, \theta) - H(\mathcal{T}_i, \theta))]\} \int_{B_{i+1}} q(\mathcal{T}_{i+1}, t_0, \theta, \theta') \lambda(d\theta') \\ &\quad + [1 - \xi \min\{1, \exp[-(H(\mathcal{T}_{i-1}, \theta) - H(\mathcal{T}_i, \theta))]\} \\ &\quad \quad - \xi \min\{1, \exp[-(H(\mathcal{T}_{i+1}, \theta) - H(\mathcal{T}_i, \theta))]\}] \int_{B_i} q(\mathcal{T}_i, t_0, \theta, \theta') \lambda(d\theta'). \end{aligned}$$

For  $i = 1$ , the above equation will be modified to the following form:

$$\begin{aligned} P((\mathcal{T}_1, \theta), A) &= \xi \min\{1, \exp[-(H(\mathcal{T}_2, \theta) - H(\mathcal{T}_1, \theta))]\} \int_{B_2} q(\mathcal{T}_2, t_0, \theta, \theta') \lambda(d\theta') \\ &\quad + [1 - \xi \min\{1, \exp[-(H(\mathcal{T}_2, \theta) - H(\mathcal{T}_1, \theta))]\}] \int_{B_1} q(\mathcal{T}_1, t_0, \theta, \theta') \lambda(d\theta'). \end{aligned}$$

Similarly, for  $i = N$  we have

$$\begin{aligned} P((\mathcal{T}_N, \theta), A) &= \xi \min\{1, \exp[-(H(\mathcal{T}_{N-1}, \theta) - H(\mathcal{T}_N, \theta))]\} \int_{B_{N-1}} q(\mathcal{T}_{N-1}, t_0, \theta, \theta') \lambda(d\theta') \\ &\quad + [1 - \xi \min\{1, \exp[-(H(\mathcal{T}_{N-1}, \theta) - H(\mathcal{T}_N, \theta))]\}] \int_{B_N} q(\mathcal{T}_N, t_0, \theta, \theta') \lambda(d\theta'). \end{aligned}$$

In order to show that the measure  $\pi$  with density  $g$  is invariant for the chain  $X_k$ , we have to prove that the following equation holds for any  $A \in \mathcal{S}$ :

$$\pi(A) = \sum_{i=1}^N \int_{\overline{\Theta}} g(\mathcal{T}_i, \theta) P((\mathcal{T}_i, \theta), A) \lambda(d\theta). \quad (4.4)$$

Suppose,  $d = \sum_{i=1}^N \int_{\overline{\Theta}} \exp(-H(\mathcal{T}_i, \theta')) \lambda(d\theta')$ . Then  $g(\mathcal{T}_i, \theta) = \frac{1}{d} \exp(-H(\mathcal{T}_i, \theta))$ .

For  $1 \leq i \leq N$  and fixed  $\theta$ , we denote  $\int_{B_i} q(\mathcal{T}_i, t_0, \theta, \theta') \lambda(d\theta')$  by  $v_i$  and for  $2 \leq i \leq N$  and fixed  $\theta$ , we define  $u_i$  as

$$u_i \stackrel{\text{def}}{=} \frac{\xi}{d} \min\{\exp(-H(\mathcal{T}_i, \theta)), \exp(-H(\mathcal{T}_{i-1}, \theta))\} (v_{i-1} - v_i).$$

Then, for  $2 \leq i \leq N - 1$ ,

$$\begin{aligned}
g(\mathcal{T}_i, \theta)P((\mathcal{T}_i, \theta), A) &= \frac{\xi}{d} \min\{\exp(-H(\mathcal{T}_i, \theta)), \exp(-H(\mathcal{T}_{i-1}, \theta))\} \cdot v_{i-1} \\
&\quad + \frac{\xi}{d} \min\{\exp(-H(\mathcal{T}_i, \theta)), \exp(-H(\mathcal{T}_{i+1}, \theta))\} \cdot v_{i+1} \\
&\quad + \left[\frac{1}{d} \exp(-H(\mathcal{T}_i, \theta)) - \frac{\xi}{d} \min\{\exp(-H(\mathcal{T}_i, \theta)), \exp(-H(\mathcal{T}_{i-1}, \theta))\}\right. \\
&\quad \quad \left. - \frac{\xi}{d} \min\{\exp(-H(\mathcal{T}_i, \theta)), \exp(-H(\mathcal{T}_{i+1}, \theta))\}\right] \cdot v_i \\
&= u_i - u_{i+1} + \frac{1}{d} \exp(-H(\mathcal{T}_i, \theta)) \cdot v_i.
\end{aligned} \tag{4.5}$$

Similarly, we have

$$g(\mathcal{T}_1, \theta)P((\mathcal{T}_1, \theta), A) = -u_2 + \frac{1}{d} \exp(-H(\mathcal{T}_1, \theta)) \cdot v_1$$

and

$$g(\mathcal{T}_N, \theta)P((\mathcal{T}_N, \theta), A) = u_N + \frac{1}{d} \exp(-H(\mathcal{T}_N, \theta)) \cdot v_N.$$

Since  $\sum_{i=2}^{N-1} (u_i - u_{i+1}) = u_2 - u_N$ , by its telescopic nature, we have

$$\begin{aligned}
&\sum_{i=1}^N g(\mathcal{T}_i, \theta)P((\mathcal{T}_i, \theta), A) \\
&= \sum_{i=2}^{N-1} \left[ (u_i - u_{i-1}) + \frac{1}{d} \exp(-H(\mathcal{T}_i, \theta)) \cdot v_i \right] - u_2 \\
&\quad + \frac{1}{d} \exp(-H(\mathcal{T}_1, \theta)) \cdot v_1 + u_N + \frac{1}{d} \exp(-H(\mathcal{T}_N, \theta)) \cdot v_N \\
&= \sum_{i=1}^N \frac{1}{d} \exp(-H(\mathcal{T}_i, \theta)) \cdot v_i \\
&= \sum_{i=1}^N \frac{1}{d} \exp(-H(\mathcal{T}_i, \theta)) \cdot \int_{B_i} q(\mathcal{T}_i, t_0, \theta, \theta') \lambda(d\theta').
\end{aligned} \tag{4.6}$$

For a fixed  $\mathcal{T}_i$ , the diffusion process obtained in the diffusion step of our algorithm has the invariant measure whose density with respect to  $\lambda$  is given by  $\frac{1}{c} \exp(-H(\mathcal{T}_i, \theta))$ , where  $c = \int_{\Theta} \exp(-H(\mathcal{T}_i, \theta)) \lambda(d\theta)$  (see Keenan and Shorter [12]). Therefore, if we sample from that process at a fixed time interval  $t_0$ , the resulting Markov chain will have the same invariant measure and

$$\int_{\Theta} \exp(-H(\mathcal{T}_i, \theta)) q(\mathcal{T}_i, t_0, \theta, \theta') \lambda(d\theta) = \exp(-H(\mathcal{T}_i, \theta')). \tag{4.7}$$

From (4.6),

$$\begin{aligned}
& \sum_{i=1}^N \int_{\Theta} g(\mathcal{T}_i, \theta) P((\mathcal{T}_i, \theta), A) \lambda(d\theta) \\
&= \sum_{i=1}^N \frac{1}{d} \int_{B_i} \int_{\Theta} \exp(-H(\mathcal{T}_i, \theta)) q(\mathcal{T}_i, t_0, \theta, \theta') \lambda(d\theta) \lambda(d\theta') \\
&= \sum_{i=1}^N \frac{1}{d} \int_{B_i} \exp(-H(\mathcal{T}_i, \theta')) \lambda(d\theta') \quad [\text{using (4.7)}] \\
&= \sum_{i=1}^N \int_{B_i} g(\mathcal{T}_i, \theta) \lambda(d\theta) \\
&= \pi(A).
\end{aligned}$$

Hence, we have proved that (4.4) holds for any  $A \in \mathcal{S}$  and that proves  $\pi$  is an invariant measure for the Markov chain  $X_k$ . Clearly,  $\pi$  is a probability measure.  $\square$

**THEOREM 4.13.** The chain  $X_k$  is an ergodic chain with the stationary probability measure  $\pi$ , i.e., for any probability measure  $\mu$  on  $\mathcal{S}$ ,

$$\lim_{n \rightarrow \infty} \|\mu P_n - \pi\| = 0.$$

*Proof.* In Theorem 4.11 we have proved that the Markov chain  $X_k$  is aperiodic. Hence (4.2) holds for any pair of probability measures  $(\nu_1, \nu_2)$  on  $\mathcal{S}$ . According to Theorem 4.12,  $\pi$  is an invariant probability measure for  $P$ . So  $\pi P = \pi$  and hence  $\pi P_n = \pi$ . Now take  $\nu_1 = \mu$  and  $\nu_2 = \pi$ .  $\square$

Theorem 4.13 tells us that, under this transition described by our algorithm, any probability measure evolves to a stationary measure whose density is given by (4.3).

**5. Application of the methodology.** Consider the four time-sampled concentration profiles displayed in Figure 1, depicting two fundamental biological phenomena: disease (left) and aging (right). The concentrations are those of two pituitary-released hormones: adrenocorticotropin hormone (ACTH) from both a healthy individual and a subject with Cushing's disease (a pituitary tumor), and luteinizing hormone (LH) from both pre- and postmenopausal females. It is ACTH which regulates the levels of cortisol which are produced and secreted by the adrenal gland, and it is LH which regulates the levels of estrogen and progesterone resulting or not resulting in ovulation. The extraction from these concentrations of their unobserved secretion and elimination rates is a necessary first step before clinical interpretation and decision-making. Cushing's disease occurs when a pituitary tumor develops, and the transformed ACTH-secreting cells not only secrete excessive levels of ACTH, but they are unresponsive to the resulting elevated cortisol levels. An immediate question is whether, in such a situation, an algorithm could (automatically) detect that an inappropriate amount of ACTH was being secreted. Also in Figure 1 are LH profiles for pre- and post-menopausal females. The premenopausal LH was obtained during the beginning (early follicular) phase of the menstrual cycle, where the levels of estrogen and progesterone are comparable to those of a postmenopausal



female. Hence, the LH differences represent the effects of aging. An interesting second question is whether an algorithm, based upon observed levels of LH concentration, could assess how close a premenopausal female is to menopause (i.e., calculate a “biological age”). In Figures 2-3 are displayed, for each of the four concentration profiles, the  $\mathbb{T}$  (the collections of putative pulse sets) and the surfaces from which they were constructed (Section 3.1).

Let  $\hat{s} = (\mathcal{T}_{\hat{m}}, \hat{\theta})$  be a realization from the posterior measure on  $S = \mathbb{T} \times \overline{\Theta}$ , so that there are  $\hat{m}$  pulse times:  $\mathcal{T}_{\hat{m}}$  and

$$\hat{\theta} = (\hat{\beta}_0, \hat{\alpha}_1, \hat{\alpha}_2, \hat{\eta}_1, \hat{\eta}_2, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3, \hat{\sigma}_A, \hat{\sigma}_\epsilon).$$

For this value, one can then calculate the conditional expectation of the unobserved secretion rate  $Z(\cdot)$ , based upon the data  $Y$ :

$$\hat{Z}(t_i) \stackrel{\text{def}}{=} E_{\hat{s}}(Z(t_i)|Y) = \hat{\beta}_0 + \sum_{T_j \leq t_i} [\hat{\eta}_0 + \hat{\eta}_1 \times (T_j - T_{j-1}) + \hat{A}_j] \hat{\psi}(t_i - T_j), \quad \text{for } t_i, i = 1, \dots, n,$$

where one first calculates:  $\hat{A}_j = E_{\hat{s}}(A_j|Y)$ , the “predicted” value of the random component of the  $j$ -th pulse mass  $M_j, j = 1, \dots, \hat{m}$ , and where  $\hat{\psi}$  is

$$\hat{\psi}(r) = \frac{\hat{\beta}_3}{\Gamma(\hat{\beta}_1) \hat{\beta}_2^{(\hat{\beta}_1 \hat{\beta}_3)}} r^{(\hat{\beta}_1 \hat{\beta}_3) - 1} e^{-(r/\hat{\beta}_2)^{\hat{\beta}_3}}.$$

From the estimated elimination rates  $\hat{\alpha}_1, \hat{\alpha}_2$ , one can then reconvolve  $\hat{Z}$  with a biexponential elimination rate and obtain the corresponding “fit” to the data:

$$\hat{X}(t) = (a^{(1)} e^{-\hat{\alpha}_1 t} + a^{(2)} e^{-\hat{\alpha}_2 t}) X(0) + \int_0^t (a^{(1)} e^{-\hat{\alpha}_1(t-s)} + a^{(2)} e^{-\hat{\alpha}_2(t-s)}) \hat{Z}(s) ds.$$

In Figures 4-5 are displayed the results of this methodology for the ACTH and LH data, respectively. For each of these, in the top panels are displayed the resulting 100 realizations, respectively, of  $\hat{X}(\cdot), \hat{Z}(\cdot), \hat{\psi}(\cdot)$ . Then, in the bottom panels are probability distributions, based upon the 100 realizations, for various statistics of interest: # pulses, half-lives of elimination, 24-hr total basal and total pulsatile secretions, and mass per pulse. Also, there is a circadian rhythm influence on ACTH (via the hypothalamus), with stronger ACTH secretion at night. In Cushing’s disease, the circadian rhythm will be significantly lost; a measure of this was included in the estimation (a day/night change in the  $\psi$  function). From such probability distributions, one can then calculate the probability statements described in the Introduction. The results are summarized in Table 1. Interestingly, Cushing’s disease doesn’t seem to induce a change in the pulse frequency, but mainly in pulse mass. Menopause, on the other hand, appears to induce an increase in both frequency and mass. Neither situation appears to affect the rates of elimination. An additional application appears in Keenan, Chattopadhyay and Veldhuis [10] and Chattopadhyay [3].

**6. Summary.** The brain signals to the pituitary gland intermittently, thereby governing the secretion of blood-borne hormones that regulate growth, metabolism, reproduction and stress-related adaptations throughout the human lifespan. Development,

TABLE 1. ACTH and LH Posterior Distributions Summary Statistics

## ACTH - Quantiles (.05, .25, .5, .75, .95) and (Mean, SD)

	ACTH (Healthy)	ACTH (Cushing's Disease)
Slow half-life (min)	8, 22, 24, 25, 25 (21.6, 5.6)	13, 17, 18, 19, 21 (18, 2)
Total basal secretion (pg/mL)	0, 0, 14, 81, 631 (106, 199)	3957, 4776, 5561, 6229, 7394 (5525, 1024)
Total secretion (pg/mL)	1106, 1130, 1163, 1283, 2960 (1406, 568)	7006, 7502, 7789, 8360, 10004 (8015, 838)
Number of pulses	19, 19, 19, 25, 28 (22, 4)	19, 19, 24, 28, 32 (25, 4)
Mass per pulse (pg/mL)	47, 53, 67, 74, 154 (74, 32)	265.1, 296, 371, 420, 478 (365, 74)
Change point (min)	1213, 1229, 1238, 1248, 1270 (1237, 23)	1248.7, 1429, 1443, 1450, 1450 (1417, 87)

## LH - Quantiles (.05, .25, .5, .75, .95) and (Mean, SD)

	LH (Pre-M)	LH (Post-M)
Slow half-life (min)	46, 52, 54, 56, 60 (54, 4)	109, 117, 120, 125, 148 (122, 10)
Total basal secretion (IU/L)	14, 28, 34, 36, 41 (32, 8)	161, 207, 231, 257, 314 (231, 53)
Total secretion (IU/L)	109, 116, 118, 122, 136 (119, 7)	720, 826, 856, 886, 931 (853, 54)
Number of pulses	18, 18, 18, 19, 24 (19, 2)	28, 28, 28, 28, 31 (28, 1)
Mass per pulse (IU/L)	6, 6, 7, 7, 7 (7, 1)	28, 31, 32, 33, 36 (32, 2)

aging and disease modify neuroendocrine control. However, only pituitary output is observable in the human. In the present paper, a methodology was developed and justified by which one can make probability statements about the underlying pulsing, secretion and elimination parameters, based upon the pituitary output. The algorithm consisted of two stages. In the first stage, the data (a hormone concentration time-profile) is “selectively smoothed” by a nonlinear parabolic PDE whose diffusion coefficient is inversely related to the degree of rapid increase. This process generates a collection of potential pulse time sets ( $\mathbb{T}$ ). In the second stage, a prior probability measure is placed on  $\mathbb{T} \times \bar{\Theta}$ , and via an algorithm which alternates between a Metropolis algorithm on  $\mathbb{T}$  and a time-homogeneous diffusion on  $\bar{\Theta}$ , simulation from the posterior measure is achieved. It was then applied to ACTH and LH data.

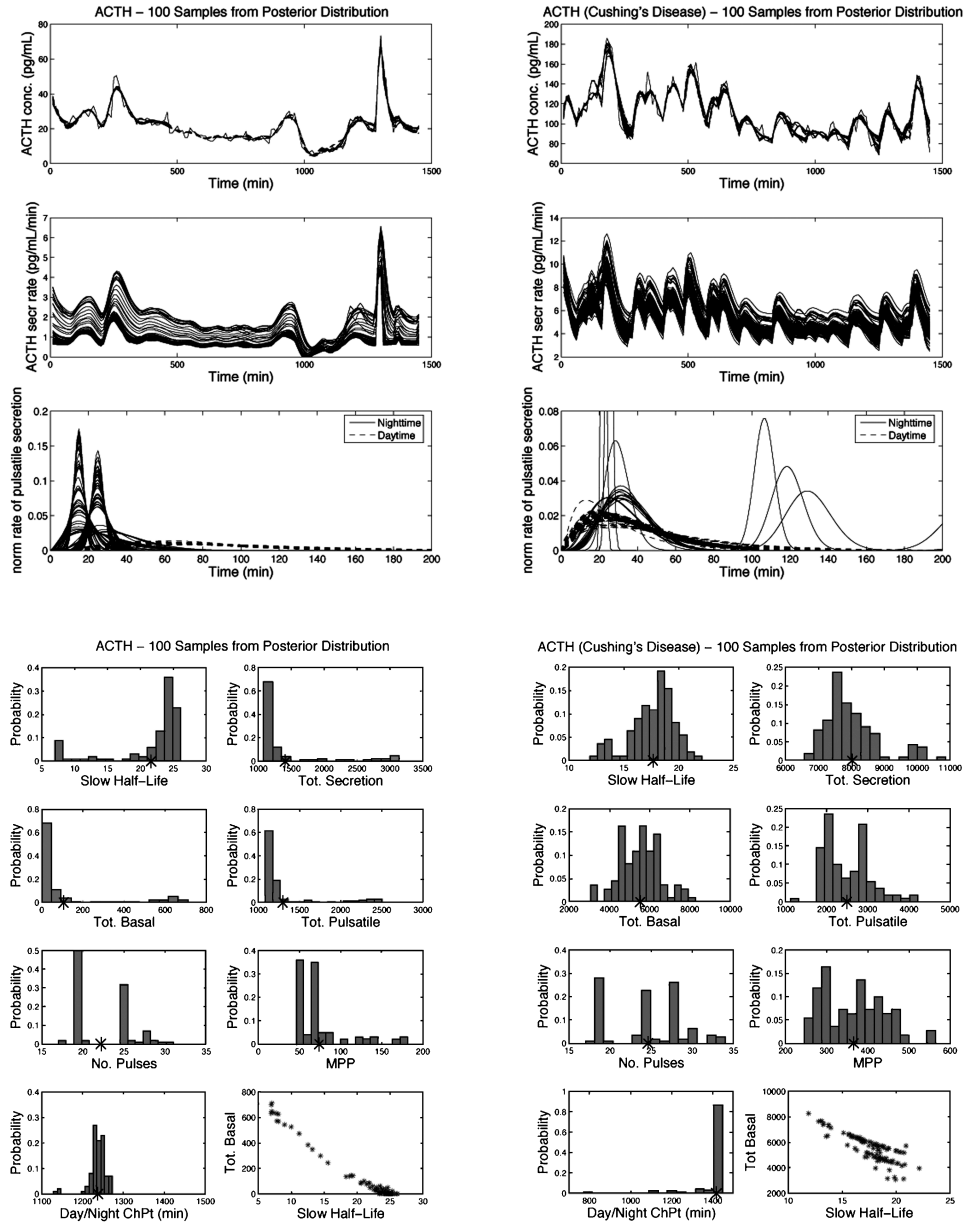


FIG. 4. Top: For ACTH, healthy subject (left) and Cushing's disease (right), 100 samples from the posterior distribution, based upon the data. The resulting fits, secretion rates, and the rates of release are displayed. Bottom: The estimated posterior distributions of various parameters.

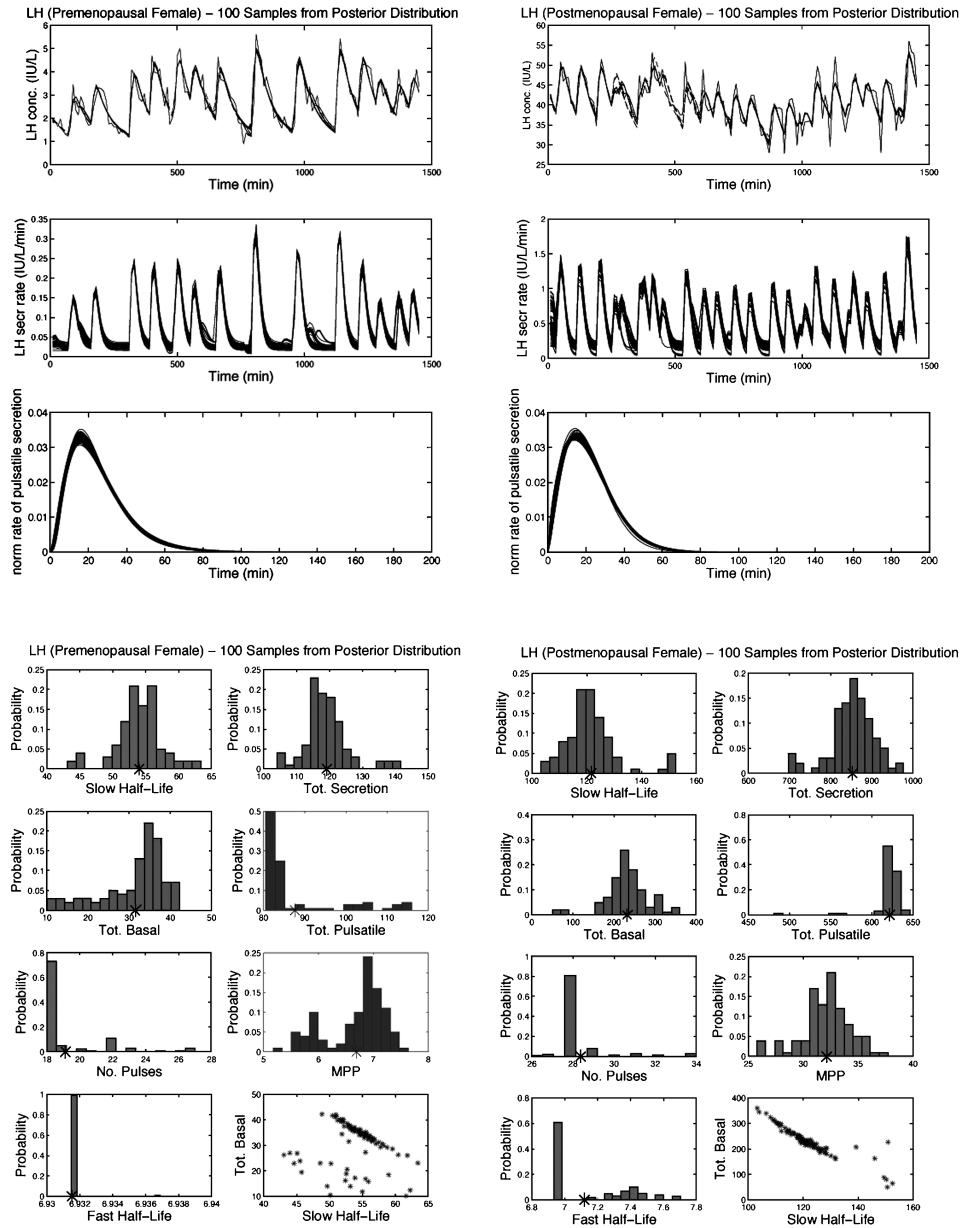


FIG. 5. Top: For LH, Premenopausal (left) and Postmenopausal (right), 100 samples from the Posterior distribution, based upon the data. The resulting fits, secretion rates, and the rates of release are displayed. Bottom: The estimated posterior distributions of various parameters.

## REFERENCES

- [1] Alvarez, L., Lions, P.-L. and Morel, J.-M. (1992). Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM J. Numer. Anal.* **29** 845-866. MR1163360 (93a:35072)
- [2] Boothby, W. M. (1975). An introduction to differentiable manifolds and Riemannian geometry. Pure and Applied Mathematics **63**, Academic Press, New York. MR0426007 (54:13956)
- [3] Chattopadhyay, S. (2001). Simultaneous Hormone Pulse Time and Secretion/Elimination Estimation: An Alternating Metropolis and Diffusion Scheme. Ph.D. thesis, Department of Statistics, University of Virginia, Charlottesville, Virginia.
- [4] Geman, S. and Hwang, C.-R. (1986). Diffusions for global optimization. *SIAM J. Control and Optimization* **24** 1031-1043. MR0854068 (87j:49064)
- [5] Grenander, U. (1983). *Tutorial in Pattern Theory*. Division of Applied mathematics, Brown University, Providence, Rhode Island.
- [6] Grenander, U. and Miller, M. (1994). Representations of knowledge in complex systems. *J. Royal Statist. Soc. Ser B*, **56**, 549-581 (Discussion, 581-603). MR1293234 (96e:62003)
- [7] Ito, S. (1957). Fundamental solutions of parabolic differential equations and boundary value problems. *Japan J. Math.* **27** 55-102. MR0098240 (20:4702)
- [8] John, F. (1982). *Partial Differential Equations*. Fourth Edition. Springer-Verlag, New York. MR0831655 (87g:35002)
- [9] Keenan, D.M., Alexander, S.L., Irvine, C.H.G., Clarke, I.J., Canny, B.J., Scott, C.J., Tilbrook, A.J., Turner, A.I., Veldhuis, J.D. (2004). Reconstruction of in vivo time-evolving neuroendocrine dose-response properties unveils admixed deterministic and stochastic elements in interglandular signaling. *Proc. Natl. Acad. Sciences* **101** 6740-6745.
- [10] Keenan, D. M., Chattopadhyay, S., and Veldhuis, J. D. (2005). Composite model of time-varying appearance and disappearance of neurohormone pulse signals in blood. *J. Theoret. Biol.* **236**, 242-255. MR2157306.
- [11] Keenan, D. M., Licinio, J. and Veldhuis, J. D. (2001). A feedback-controlled ensemble model of the stress-responsive hypothalamo-pituitary-adrenal axis. *Proc. Natl. Acad. Sciences* **98** 7:4028-4033.
- [12] Keenan, D. M. and Shorter, P. A. (2004). Simulation by diffusion on a manifold with boundary: Applications to ultrasonic prenatal medical imaging. *SIAM J. Appl. Math.* **64** 3:932-960. MR2068448 (2005b:92021)
- [13] Keenan, D. M., Sun, W. and Veldhuis, J. D. (2000). A stochastic biomathematical model of the male reproductive hormone system. *SIAM J. Appl. Math.* **61** 3:934-965. MR1788025 (2001j:92016)
- [14] Keenan, D. and Veldhuis, J. D. (2001). Disruptions in the hypothalamic luteinizing-hormone pulsing mechanism in aging men. *Amer. J. Physiol.*, 281: R1917-R1924.
- [15] Keenan, D. M., Veldhuis, J. D. and Yang, R. (1998). Joint recovery of pulsatile and basal hormone secretion by stochastic nonlinear random-effects analysis. *Amer. J. Physiol.* **275**, R1939-R1949.
- [16] Mauger, D.T., Brown, M. B. and Kushler, R. H. (1995). A comparison of methods that characterize pulses in a time series. *Statistics in Medicine* **14** 311-325.
- [17] Perona, P. and Malik, J. (1987). Scale space and edge detection using anisotropic diffusion. *Proc. IEEE Computer Society Workshop on Computer Vision*.
- [18] Revuz, D. (1984). *Markov Chains*. Second edition. (*North-Holland Mathematical Library* **11**). North-Holland Publishing Co., Amsterdam. MR0415773 (54:3852)
- [19] Sato, K. and Ueno, T. (1965). Multi-dimensional diffusion and the Markov process on the boundary. *J. Math. Kyoto Univ.* **4** 529-605. MR0198547 (33:6702)
- [20] Veldhuis, J. D. (1995). Pulsatile hormone release as a window into the brain's control of the anterior pituitary gland in health and disease: implications and consequences for pulsatile luteinizing hormone secretion. In *The Endocrinologist*, Editor: L. Loriaux. Williams & Wilkins, Baltimore.
- [21] Yang, R. (1997). Maximum Likelihood Estimation Asymptotics for Parameter-Dependent Mixed Effects Models with Applications to Hormone Data. Ph.D. thesis, Division of Statistics, Department of Mathematics, University of Virginia, Charlottesville, Virginia.