A Stochastic Deconvolution Method to Reconstruct Insulin Secretion Rate After a Glucose Stimulus

Giovanni Sparacino and Claudio Cobelli,* *Member, IEEE*

Abstract— **Insulin secretion rate (ISR) is not directly measurable in man but it can be reconstructed from C-peptide (CP) concentration measurements by solving an input estimation problem by deconvolution. The major difficulties posed by the estimation of ISR after a glucose stimulus, e.g., during an intravenous glucose tolerance test (IVGTT), are the ill-conditioning of the problem, the nonstationary pattern of the secretion rate, and the nonuniform/infrequent sampling schedule. In this work, a nonparametric method based on the classic Phillips–Tikhonov regularization approach is presented. The problem of nonuniform/infrequent sampling is addressed by a novel formulation of the regularization method which allows the estimation of quasi time-continuous input profiles. The input estimation problem is stated into a Bayesian context, where the *a priori* known nonstationary characteristics of ISR after the glucose stimulus are described by a stochastic model. Deconvolution is tackled by linear minimum variance estimation, thus allowing the derivation of new statistically based regularization criteria. Finally, a Monte-Carlo strategy is implemented to assess the uncertainty of the estimated ISR arising from CP measurement error and impulse response parameters uncertainty.**

I. Introduction

THE ability to measure insulin secretion rate (ISR) is essential for a quantitative understanding of the glucose regulation system in man, both in healthy and disease states. Unfortunately, ISR cannot be directly measured since insulin is secreted by the pancreatic β-cells into the portal vein which is not accessible *in vivo*: one can only measure the effect of secretion in the circulation, i.e., the plasma concentration of insulin. Furthermore, before reaching plasma, insulin undergoes a large and variable liver extraction [40]. Thus, insulin plasma concentration only reflects post-hepatic insulin delivery rate into the circulation. Luckily, a peptide, C-peptide (CP), is co-secreted with insulin on an equimolar basis but, unlike insulin, it is not extracted by the liver [37]. Thus the plasma concentration of CP directly reflects the pancreatic ISR. Since CP kinetics are linear it is possible to pose the measurement of pancreatic ISR as an input estimation problem which can be solved by deconvolution [17], [37], [38].

The aim of this study is to quantify ISR in humans after a glucose stimulus. In particular we consider the intravenous glucose tolerance test (IVGTT), where an impulse dose of glucose is administered. This test is widely used in physiological and clinical investigations to assess glucose tolerance, see e.g., [3], [6]. The major challenges posed in the reconstruction of ISR during IVGTT by deconvolution are the ill-conditioning of the problem, the nonuniform/infrequent sampling and the nonstationary pattern of the ISR signal.

The first attempts to use deconvolution to quantify glucose-induced ISR trace back to [46] and [36]. In these studies insulin plasma concentration was used, and therefore, only post-hepatic insulin delivery rate has been reconstructed. In [46] plasma insulin values are first transformed by reading off a "smooth line" drawn by eye to pass between each pair of duplicate measurements; then, insulin delivery rate was obtained through the direct solution of the input estimation problem, since the short insulin half-life makes the deconvolution problem only slightly ill-conditioned, see e.g., [9]. In [36], post-hepatic insulin delivery rate was reconstructed by means of the classic Phillips–Tikhonov regularization method [35], [44]; the problem of nonuniform/infrequent sampling was tackled by interpolating the sampled points with straight lines and assigning the mean value of such interpolating functions to frequent and equally-spaced time intervals. More recently, investigators have resorted to CP concentration to assess ISR. In particular, in [41], CP, and thus insulin, secretion rate during an IVGTT was estimated by inverting analytically the two-compartment CP kinetic model by using splines to smooth the CP concentration and calculate numerically its first time-derivative, as proposed in [17]. This procedure assumes the CP impulse response to be biexponential. Moreover, smoothing output data before making deconvolution can be critical and it does not always guarantee against ill-conditioning, see e.g., [15]. Of note is that with all the above approaches [36], [41], [46] it is difficult to obtain confidence limits of the estimated ISR profile.

The nonparametric deconvolution approach we apply here to the same data set used in [41] is an evolution of the Phillips–Tikhonov regularization method. The problem of nonuniform/infrequent sampling is addressed by a novel formulation of the regularization method which allows the estimation of quasi time-continuous input profiles. The input estimation problem is stated into a stochastic context so that regularization is tackled by solving a linear minimum variance estimation problem where the *a priori* known nonstationary characteristics of the input are described by a stochastic model. The major advantages provided by the stochastic context consist in providing new statistically based regularization

Manuscript received August 1, 1994; revised December 19, 1995. This work was supported in part by MURST under Project "Bioingegneria dei Sistemi Metabolici e Cellulari" and by National Institutes of Health under Grant RR-02176, "Resource Facility for Kinetic Analysis." *Asterisk indicates corresponding author.*

G. Sparacino is with the Dipartimento di Elettronica ed Informatica, Università di Padova, Padova, 35100 Italy.

*C. Cobelli is with the Dipartimento di Elettronica ed Informatica, Università di Padova, Via Gradenigo 6/A, Padova 35100 Italy (e-mail: cobelli@pia.dei.unipd.it).

Publisher Item Identifier S 0018-9294(96)03187-4.

criteria and obtaining expressions to easily compute the confidence intervals of the estimate.

The paper organization is as follows. Section II states the deconvolution problem. The experiment and the data base are described in Section III. Section IV is devoted to the modeling of the CP impulse response. Section V first recalls some fundamentals about deconvolution and its ill-conditioning; then, a brief review of the Phillips–Tikhonov regularization method is presented. Its use to solve our problem is shown to be not successful, thus calling for new strategies. In Section VI, deconvolution is first restated within a stochastic context as a linear minimum variance estimation problem. Then the issues of nonuniform/infrequent sampling, input nonstationarity, and statistically based choice of the regularization parameter are addressed. Finally, the question of confidence intervals is considered. Linear minimum variance estimation naturally provides analytical expressions to compute confidence intervals. However, since these intervals assume an error free impulse response model, a Monte-Carlo strategy is used to evaluate the joint uncertainty of the estimate arising from both data error and impulse response parameters uncertainty. Finally, a discussion is reported in Section VII. Two appendices complement the paper. Appendix A shows the statistical basis of the new regularization criteria. Appendix B tests the deconvolution procedure on a simulation problem.

II. Estimation of Insulin Secretion Rate: A Deconvolution Problem

CP kinetics are known to be linear and time-invariant among a wide range of concentration levels [22], [38]. So it is possible to relate CP plasma concentration and insulin (equal to the CP) secretion rate by the convolution integral

$$c(t) = \int_{-\infty}^{t} g(t-\tau)\mathrm{ISR}(\tau)\, d\tau \tag{1}$$

where:

- $c(t)$ is the CP plasma concentration (pmol/ml).
- $g(t)$ is the CP impulse response (1/ml).
- $\mathrm{ISR}(t)$ is the insulin secretion rate (pmol/min).

Thus, the estimation of ISR during IVGTT from the sampled concentration can be stated as a deconvolution problem. To solve it, the impulse response $g(t)$ of the CP system is needed. Usually the protocol to determine the impulse response (see Section III) is performed separately from the IVGTT. In theory, there could thus be the risk that the so-determined $g(t)$ is different from that during the IVGTT. Available studies support, however, that $g(t)$ is not affected by elevated values of glucose and insulin concentrations like those observed during an IVGTT [31] and that intraindividual variability is small, see e.g., [38] and [49]. We thus assume the impulse response not to vary between the two experiments, also in keeping with [41] and [45].

III. Experiment and Data Base

Data base originates from a two-stage experiment already described in [41] and performed on seven young normal men. To determine in each individual the impulse response of the CP system, a somatostatin infusion was administered to suppress endogenous pancreatic secretion and was maintained throughout the study. An intravenous bolus (150 μg equivalent to 49 650 pmol) of biosynthetic CP [38] was then injected in the blood. CP plasma samples were collected at minutes: 2–11, 14, 17, 20, 25, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 110, 120, 140, 160, and 180.

The same subjects were then observed during an IVGTT. Basal condition was monitored for 1 h by measuring plasma concentration levels every 15 min at times -60, -45, -30, -15, and zero. An intravenous bolus of glucose (0.5 g/Kg of body weight) was then injected in the blood. CP concentrations were measured for four hours, collecting data at minutes: 1, 3, 5, 7, 9, 12, 15, 20, 25, 30, 35, 40, 45, 60, 75, 90, 105, 120, 140, 160, 180, 200, 220, and 240.

Measurement error of CP concentration was assumed to be uncorrelated, Gaussian with zero mean. Its coefficient of variation (CV) is assumed to be constant and ranging between 4–6%.

A representative data set of the two-stage experiment is shown in Fig. 1.

IV. The Impulse Response

A sum of exponential model describes the impulse response of the CP system, i.e., the data following the impulse dose of CP

$$g(t) = \sum_{i=1}^{M} A_i e^{-\alpha_i t}. \tag{2}$$

Parameter estimation was performed by a nonlinear least squares algorithm assuming a constant (but unknown) coefficient of variation of the measurement error. The CV has been estimated *a posteriori* [5], [30], also given its crucial role in regularized deconvolution (see Sections V and VI).

It is commonly accepted [17], [22], [37] that a two-exponential model is a good description of the CP impulse response. In this study we have found that a three-exponential model (where the third exponential is the fastest one) was generally superior to the two-exponential one by considering both the Akaike [1] and the Schwarz [42] criteria. In other words, by increasing the model-order, there is a significant decrease of the weighted residual sum of squares with only a slight concomitant deterioration of the parameter precision. For example, by considering the Akaike information criterion AIC $= 2P + N \log \mathrm{WRSS}$ [30], where P is the number of model parameters, N is the number of data points, and WRSS is the weighted (by the inverse of squared measurements) residual sum of squares, the three-exponential model was superior in all subjects except for subject #1 and subject #7, where the two-exponential model had a comparable performance.

Table I shows the parameters of the three-exponential model, the difference of AIC, ΔAIC, between the three and the two-exponential model, and the *a posteriori* estimate of the measurement error CV.

Fig. 2 shows for subjects #1–5 and #7 the fit of the three-exponential model. For subject #6, the fit was presented in

Fig. 1. The influence of the CP impulse response model order, in particular two versus three, in reconstructing various ISR patterns by deconvolution has been recently studied in [43]; the presence of high frequency components in the ISR spectrum during IVGTT encourages the use of the three-exponential model.

V. Deconvolution: Fundamentals and Review of the Regularization Approach

A. Problem Statement

In general, deconvolution techniques consider problems involving causal signals, i.e., the system initial conditions are assumed to be zero. In our case, since for $t < 0$ a basal secretion occurs, it is necessary to remove its effect before deconvolving the data, thus obtaining a system with initial conditions equal to zero. From (1) it follows

$$c(t) = \int_{-\infty}^{0} g(t-\tau)\mathrm{ISR}(\tau)\,d\tau + \int_{0}^{t} g(t-\tau)\mathrm{ISR}(\tau)\,d\tau \quad (3)$$

where time zero is the time of glucose injection. Assuming ISR to be constant at level ISR_b for $t < 0$, its effect for $t > 0$, say $c_0(t)$, is given by

$$c_0(t) = \int_{-\infty}^{0} g(t-\tau)\mathrm{ISR}_b\,d\tau. \quad (4)$$

ISR_b can be estimated by multiplying the basal CP concentration by the clearance rate estimated from the impulse response [5]. Basal concentration is determined as the mean of the five samples collected before the glucose injection (see Section III). Letting

$$\begin{aligned} c'(t) &= c(t) - c_0(t) \\ &= \int_{0}^{t} g(t-\tau)\mathrm{ISR}(\tau)\,d\tau \end{aligned} \quad (5)$$

it is then possible to estimate $\mathrm{ISR}(t)$ from $c'(t)$ data by deconvolution.

Since c_0 is known with high precision, the error on c_0 can be assumed negligible with respect to the error on c; thus, the error on c' can be considered approximately equal to the error on c.

In conclusion, the inverse problem to solve is

$$c'(t) = \int_{0}^{t} g(t-\tau)\mathrm{ISR}(\tau)\,d\tau. \quad (6)$$

Let $\Omega_s = \{t_1, \cdots, t_k, \cdots, t_n\}$ be the given (nonuniform) sampling schedule and assume that the signal $\mathrm{ISR}(t)$ can be approximated as a piecewise constant within each sampling interval. For $t_k \in \Omega_s$, $k = 1, 2, \cdots, n$

$$\begin{aligned} c'(t_k) &= \int_{0}^{t_k} g(t_k-\tau)\mathrm{ISR}(\tau)\,d\tau \\ &= \sum_{i=1}^{k} \int_{t_{i-1}}^{t_i} g(t_k-\tau)\mathrm{ISR}(\tau)\,d\tau \\ &= \sum_{i=1}^{k} \mathrm{ISR}(t_i) \int_{t_{i-1}}^{t_i} g(t_k-\tau)\,d\tau \end{aligned} \quad (7)$$

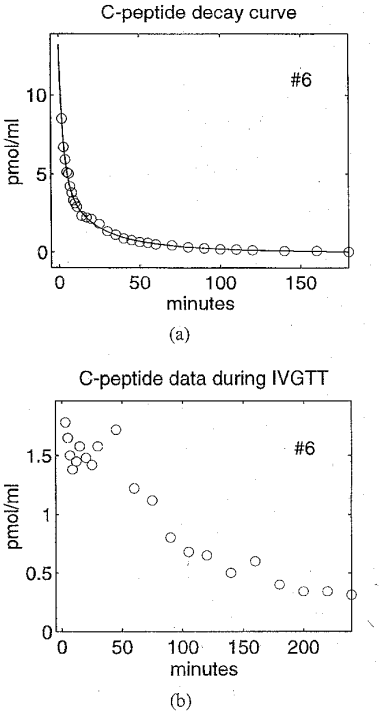


Fig. 1. The two-stage experiment performed on seven young normal subjects. A representative data set of C-peptide concentration (subject #6) is shown. (a) Following a C-peptide bolus injection. The continuous line denotes the fitted impulse response model. (b) Following a glucose bolus injection during an intravenous glucose tolerance test (IVGTT).

where $t_0 = 0$. Letting

$$g_{k,i} = \int_{t_{i-1}}^{t_i} g(t_k-\tau)\,d\tau \quad (8)$$

$$u_i = \mathrm{ISR}(t_i) \quad (9)$$

it follows

$$c'(t_k) = \sum_{i=1}^{k} u_i g_{k,i}. \quad (10)$$

One may also think u_k to be the mean level of ISR during the kth sampling interval. Adopting a matrix notation

$$c' = Gu \quad (11)$$

where c' is the n-dimensional vector of the sampled output (noise-free), u is the n-dimensional vector whose components are samples of the unknown ISR, and G is a $n \times n$ lower-triangular matrix, whose entries are

$$G(k, i) = g_{k,i} \quad k \le i. \quad (12)$$

Note that, as sampling is not uniform, relation (10) does not represent a discrete convolution, i.e., the discretized system is time-varying. $G(k, i)$ depends on both the arguments, and not on their difference only, so that G is not a Toeplitz matrix.

TABLE I
ESTIMATES OF THE THREE-EXPONENTIAL MODEL PARAMETERS (A_i, α_i = 1, 2, 3) WITH THEIR PRECISION (EXPRESSED AS PERCENT CV, IN PARENTHESIS); VALUES OF THE DIFFERENCE OF AIC, ΔAIC, BETWEEN THE THREE AND THE TWO-EXPONENTIAL MODEL, AND *A POSTERIORI* ESTIMATE OF THE MEASUREMENT ERROR CV

Subject	$A_1 \times 10^4$ $(ml)^{-1}$	$A_2 \times 10^4$ $(ml)^{-1}$	$A_3 \times 10^4$ $(ml)^{-1}$	α_1 $(min)^{-1}$	α_2 $(min)^{-1}$	α_3 $(min)^{-1}$	ΔAIC	*CV* (%)
# 1	1.28 (48)	1.02 (72)	0.495 (10)	0.281 (56)	0.104 (36)	0.026 (3)	0	5.4
# 2	1.71 (18)	0.963 (34)	0.519 (14)	0.287 (30)	0.081 (25)	0.023 (5)	-10	5.5
# 3	2.21 (12)	1.000 (8)	0.370 (18)	0.326 (16)	0.053 (13)	0.016 (7)	-33	4.8
# 4	2.23 (13)	0.946 (23)	0.531 (11)	0.339 (23)	0.079 (20)	0.022 (4)	-18	4.8
# 5	2.10 (26)	0.835 (26)	0.559 (15)	0.378 (36)	0.068 (30)	0.017 (7)	-11	7.5
# 6	1.61 (20)	0.630 (26)	0.390 (19)	0.367 (31)	0.076 (29)	0.023 (8)	-8	6.7
# 7	1.29 (39)	0.882 (66)	0.819 (9)	0.253 (49)	0.085 (40)	0.024 (3)	2	4.5

Assuming the measurement error to be additive, we model the observations as

$$\begin{aligned} y &= c' + v \\ &= Gu + v \end{aligned} \tag{13}$$

where y denotes the n-dimension vector of the noisy data and v is the n-dimension vector of the measurement errors. Let W denote the $n \times n$ covariance matrix of the measurement error; since errors are uncorrelated, W is diagonal. Assuming measurement errors to have a constant CV, the ith element of the diagonal of W, i.e., the variance of the error on the measurement taken at time t_i, can be expressed as $(\mathrm{CV}y_i)^2$ (CV here is a real number).

B. Deconvolution by Least Squares and Ill-Conditioning

Least Squares (LS) estimation is the simplest and most direct way to solve the input estimation problem of model (13). The estimate is the solution of the optimization problem

$$\min_u (y - Gu)^T W^{-1} (y - Gu). \tag{14}$$

Since G is invertible ($g_{i,i} \neq 0$, $\forall i$), LS estimate is $\hat{u} = G^{-1}y$. Residuals are therefore, zero. LS estimate is unbiased and the covariance matrix of the estimation error, $e = u - \hat{u}$, is equal to $(G^T W^{-1} G)^{-1}$.

It is well known that such a direct solution of the deconvolution problem can be severely ill-conditioned, i.e., a small percentage error in the data can be amplified in a much larger percentage error in the estimate. Ill-conditioning is enhanced by the smoothness of the impulse response as well as by a frequent sampling rate. Details can be found in [21] and [29]. An example of LS estimation applied to our case is shown in Fig. 3, where the deconvoluted profile is presented together with the model predictions (reconvolution) against the data. All the data are fully matched at the expense of large oscillations. In particular, wide oscillations during the second-phase secretion, say for $t > 10 - 15$ min, are physiologically unplausible and can be considered to be mostly due to ill-conditioning.

The ill-conditioning of the deconvolution problem is a classic in the mathematics/physics/engineering literature. Several methods attack it by assuming the analytic expression of the input to be known except for a small number of parameters so that the deconvolution problem becomes a parameter estimation problem. This approach is often named *parametric deconvolution.* In [51], for example, the unknown input is assumed to be described by a sum of exponentials and one has only to estimate the amplitudes and the eigenvalues of the exponentials. Other cases are discussed in [11], [48], [50], and [53]. In parametric deconvolution the ill-conditioning is circumvented by constraining *a priori* the functional form of the input. However, this requires to deal with issues such as the choice of the model order, e.g., the order of the polynomials in [11], and the problem of local minima in parameter estimation phase. Furthermore, the use of parametric methods makes it difficult to provide confidence limits. Even if parametric deconvolution has been successfully used in several applications, e.g., [23] and [50], the adoption of a specific functional form can be a rather heavy assumption. In several cases it is even not possible to assign a prescribed functional form to the input.

Recently, *splines* approximation of the unknown input has been proposed [54], [55]. Splines functions are a very flexible tool, see e.g., [57], and, in a deconvolution context, they can be constrained in several ways to account for possible sources of *a priori* knowledge about the unknown input (e.g., nonnegativity, monotonicity) [55]. However, as illustrated in [54], their use requires to cope with sometimes critical issues, such as the choice of the number and location of the splines knots which determine the input smoothing degree and the behavior of the estimate in presence of fast transients.

An alternative approach does not postulate any functional form of the input and it is often indicated as *nonparametric deconvolution.* Most of the nonparametric methods are based on the so-called Phillips–Tikhonov regularization method [35], [44] and start from the discretization of the convolution integral, facing ill-conditioning by applying regularization/filtering techniques based on some *a priori* information, e.g., the input

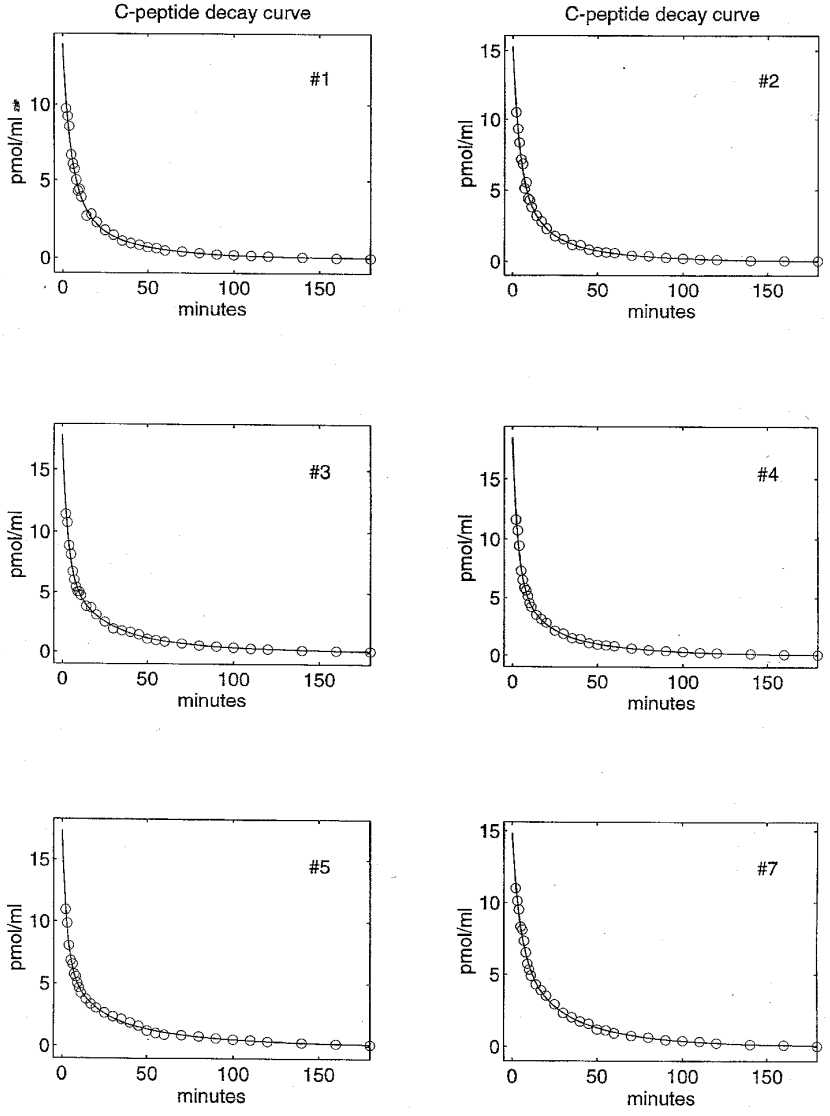


Fig. 2. Fit of the C-peptide three-exponential impulse response model (subjects #1–5 and #7).

is regular; to mention but a few contributions, we remind [10], [13], [25], [28], and [47]. For a survey, see [4]. Some nonparametric methods, in place of regularization, exploit the maximum entropy principle, e.g., [7], [8], and [16]. Below we will consider the Phillips–Tikhonov regularization method for solving our problem.

C. The Phillips–Tikhonov Regularization Approach

The idea of the Phillips–Tikhonov regularization method is to look for a solution which provides a good data fit and has, at the same time, a certain degree of smoothing. The approach assumes model (13) and estimates the unknown input u by solving the optimization problem [32]

$$\min_u (y - Gu)^T W^{-1} (y - Gu) + \gamma u^T Q^T Q u \qquad (15)$$

whose solution is

$$\hat{u} = (G^T W^{-1} G + \gamma Q^T Q)^{-1} G^T W^{-1} y. \qquad (16)$$

The cost function in (15) is made up of two terms. The first penalizes the distance between the model predictions Gu (i.e., reconvolution) and the data, i.e., how well the estimated input can match the data. The presence of matrix W ensures that the adherence to each datum is pursued according to its reliability. The second term penalizes the roughness of the estimated input through a suitable design of the smoothing matrix Q. The relative benefit of solution regularity and data fit is given by the amplitude of the parameter γ. By raising γ, the cost of roughness increases and the data match becomes relatively less important. For such a reason γ is usually called *regularization parameter*. The choice of the regularization parameter is commonly recognized as a critical problem. Several criteria

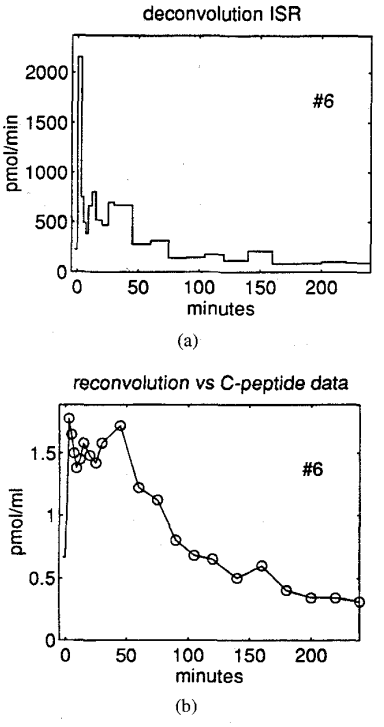


Fig. 3. Least squares approach. (a) Deconvoluted insulin secretion rate (ISR). (b) Model predictions against data (subject #6).

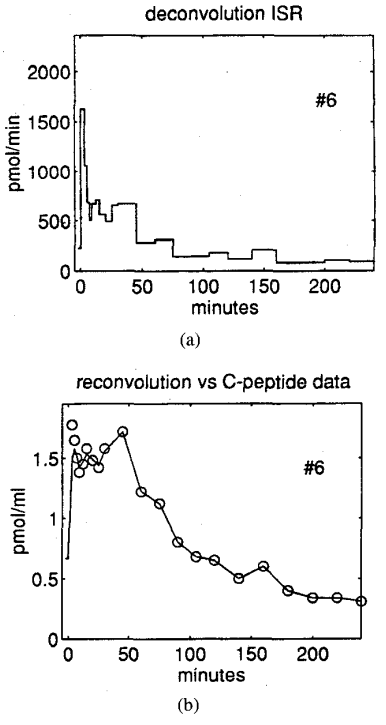


Fig. 4. The Phillips–Tikhonov approach. (a) Deconvoluted ISR. (b) Model predictions against data (subject #6).

have been proposed. When the measurement error variance is known, a widely used criterion is [47]

Criterion 1: Adjust γ until the residual sum of squares (RSS) equals the sum of the measurement error variances, i.e., until RSS = trace(W).

The rationale of the criterion is clear: since the residuals can be viewed as an *a posteriori* estimate of the measurement error (assuming an error-free model), the regularization parameter should make the residual sum of squares, RSS, equal to the expected sum of the squared measurement errors, $E[v^T v] =$ trace (W).

The smoothing addendum $u^T Q^T Q u$ must ensure the regularity of the solution. Several functionals can be selected to penalize the signal roughness [32], [44], [47] and the "most appropriate" one usually depends on the particular case under study. Usually, smoothness is pursued by penalizing the energy of the mth order time-derivatives, m being a parameter. For example, in [35] the second-order timederivatives were originally considered whereas in [10] the energy of the first derivatives was penalized so that, for uniform sampling, Q is a square lower triangular Toeplitz matrix (size n) whose first column is $[1, -2, 1, 0, \cdots, 0]^T$ or $[1, -1, 0, \cdots, 0]^T$, respectively. The parameter m is usually adjusted by trials. When sampling is nonuniform, matrix Q has to be suitably chosen since input variability must also be penalized according to the length of the time interval where it occurs. In the case of the first derivatives, for example, one can define Q by dividing by $\sqrt{\Delta t_k}$ the kth row of the above mentioned square Toeplitz matrix, Δt_k being the duration of the kth sampling interval.

The above formulation of the Phillips–Tikhonov regularization method has been applied to our problem by adopting Criterion 1 for the choice of the regularization parameter and assuming for the measurement error variance that reported in Table I. The energy of the first-order time derivatives was penalized (by considering the second order similar results were obtained). Fig. 4 shows the deconvoluted ISR and the model predictions against data for the same subject (#6) as in Fig. 3. In all subjects, the sudden ISR variations occurring immediately after the glucose injection were systematically oversmoothed. In fact, the reconvoluted profiles are not able to fit the initial rapidly changing data, say for $t < 10$–15 min, thus resulting in large residuals, see e.g., Fig. 4(b). In practice the initial residuals (one to four, depending on the subject) determine the entire RSS of Criterion 1. On the contrary, a perfect adherence to the data was obtained in the second phase, say for $t > 10$–15 min, at the expense of large oscillations (Fig. 4). A large measurement error is predicted for the initial data and virtually no error for all the remaining ones. To regularize the profile in the second-phase release, γ should be increased but this would result in an even more oversmoothing of the first-phase ISR. To obtain a good fit of the initial data, γ should be decreased but this would produce even more roughness in the second-phase ISR.

The systematic oversmoothing of the initial peak is due to the fact that the solution minimizes a cost function where the temporal variations of the input have the same price wherever they occur. In other words, the unknown signal is assumed to be "stationary." On the contrary, during IVGTT, we expect an

initial rapid release of insulin by the pancreas, i.e., ISR changes very quickly and each sample is completely unpredictable from the previous one. A much more regular profile is expected later, when insulin release is known to occur at a slower rate. Actually, this biphasic (nonstationary) pattern represents an *a priori* available knowledge which was not exploited in the above implementation of the regularization method.

Fig. 4 also shows that the staircase approximation, i.e., ISR is constant during each sampling interval, is particularly rough when the time interval gets longer due to the less frequent sampling, i.e., from 15 min on. The assumption behind (7) becomes critical and the estimated ISR is a poor approximation of a physiological continuous signal.

Finally, let us consider the regularization parameter. Its choice is critical and, in fact, this issue is a classic of the deconvolution literature. In the above strategy all the regularization effort is played in the first phase so that in the second phase (where the signal is known to be regular) ill-conditioning persists and the used regularization criterion does not play any visible role. However, neither Criterion 1 nor any of the many others proposed in the literature, for a survey see [4], [26], [34], and [39], has become a standard also because it is difficult, in a deterministic setting, to judge their relative merits.

In the next section we will exploit the lessons learned from the implementation of the classic Phillips–Tikhonov regularization method. The approximations due to the nonuniform/infrequent sampling will be addressed by means of a different formalization of the input estimation problem. Then, provided that the *a priori* knowledge about the biphasic ISR pattern can be described by a suitable model, we will state deconvolution in a stochastic embedding as a linear minimum variance estimation problem.

VI. A Stochastic Approach

Two kinds of information are available: the known biphasic pattern of ISR is an *a priori* information while the observed data can be thought as an *a posteriori* information. The above classification suggests to state the problem into a Bayesian embedding.

A. Deconvolution as a Linear Minimum Variance Estimation Problem

Consider the model

$$y = Gu + v \tag{17}$$

where y and v are a n-dimensional (stochastic) vectors, u is a N-dimensional (stochastic) vector, and G is a $n \times N$ matrix. Assume that u and v are uncorrelated with *a priori* covariance matrix Σ_u and Σ_v, respectively. For sake of reasoning, assume Σ_u and Σ_v to be positive definite matrices and u and v to have zero mean. In a probabilistic setting, consider the *Linear Minimum Variance Estimation Problem:* "Find the estimate $\hat{u}$, linearly depending on the data vector y, such that $E[||u-\hat{u}||^2]$ is minimized, i.e., minimize the expectation E of the squared euclidean norm $||\cdot||$ of the estimation error." Under the above assumptions, the solution is found by the following optimization problem [2]:

$$\min_u (y - Gu)^T \Sigma_v^{-1} (y - Gu) + u^T \Sigma_u^{-1} u. \tag{18}$$

The cost function in (18) consists of two terms. The first term denotes, as in problem (15), the distance of the model predictions from the data, whereas the second one weights the adherence to the *a priori* knowledge on u. In explicit form, the linear minimum variance estimator of u given y is

$$\hat{u} = (G^T \Sigma_v^{-1} G + \Sigma_u^{-1})^{-1} G^T \Sigma_v^{-1} y. \tag{19}$$

The estimation error $e = u - \hat{u}$ has zero mean and covariance matrix

$$\text{cov}\,(e) = (G^T \Sigma_v^{-1} G + \Sigma_u^{-1})^{-1}. \tag{20}$$

When the vectors involved in model (17) are Gaussian, the linear estimator (19) has minimum error variance among all the estimators of u given y.

To tackle deconvolution as a linear minimum variance estimation problem, the knowledge of the *a priori* second-order statistic description, i.e., mean and covariance matrix, of both u and v is thus required. Let us consider these covariance matrices depending on a scale factor (possibly unknown), i.e., $\Sigma_u = \lambda^2 \mathrm{R}$ and $\Sigma_v = \sigma^2 B$, such that (18)–(20) become

$$\min_u (y - Gu)^T B^{-1} (y - Gu) + \frac{\sigma^2}{\lambda^2} u^T R^{-1} u \tag{21}$$

$$\hat{u} = (G^T B^{-1} G + \gamma R^{-1})^{-1} G^T B^{-1} y \tag{22}$$

$$\text{cov}\,(e) = \sigma^2 (G^T B^{-1} G + \gamma R^{-1})^{-1}. \tag{23}$$

Note the similarity between the "deterministic" (15), (16) and the "stochastic" ones (21), (22). The ratio σ^2/λ^2 corresponds to γ but here it has a precise statistical meaning. When one of the two scale factors (or both of them) is (are) unknown one can estimate it (them) according to its (their) statistical meaning, thus allowing the recall of some favorable properties, such as the minimum variance of the estimator (22) and the existence of a closed form expression of the estimation error covariance matrix (23).

Finally, it is interesting to note that in [56] and [57] a similar rationale was used to interpret "deterministic" smoothing splines as Bayes estimates. This allowed the derivation of "Bayesian confidence intervals," conceptually similar to (23), whose performance have been studied in [33], [56] and [58].

B. The Virtual Grid

In this paragraph we address the problem of infrequent sampling. Consider the time-continuous model (6). With the final aim of discretizing it, let us consider two decoupled grids. Let $\Omega_s = \{t_1, t_2, \cdots, t_k \cdots, t_n\}$ be the (experimental) *sampling* grid and $\Omega_v = \{T_1, T_2, \cdots, T_k, \cdots, T_N\}$ be a finer $(N > n)$ grid over which the unknown input will be discretized. Ω_v must contain Ω_s, but it is arbitrary and it has not an experimental counterpart. For these reasons we

call Ω_v the *virtual* grid. Let $c'_v(T_k)$ denote the (noise-free) output at virtual sampling times T_k. Assume that ISR(t) can be approximated as a piecewise constant within each time interval of the virtual grid. In keeping with (7), it follows:

$$c'_v(T_k) = \sum_{i=1}^{k} \text{ISR}(T_i) \int_{T_{i-1}}^{T_i} g(T_k - \tau)\, d\tau \qquad (24)$$

where $T_0 = 0$. Adopting the usual matrix notation one has

$$c'_v = G_v u \qquad (25)$$

where c'_v and u are the vectors of input and output considered on the virtual grid, and G_v is a $N \times N$ lower triangular matrix whose entries are

$$G_v(k, i) = \int_{T_{i-1}}^{T_i} g(T_k - \tau)\, d\tau \quad k \leq i. \qquad (26)$$

Times belonging to the virtual grid Ω_v but not present in the sampling grid Ω_s do not correspond to sampled output data. We can regard them as (virtually) missing data. Denote by G the $n \times N$ matrix obtained by canceling from G_v those rows which don't correspond to sampled output data. Finally, by considering the measurement error we have a model of the observations as in (17).

When $N > n$, model (17) determines a linear system with less equations than unknowns. Indeed, the reconstruction of the N-dimension vector u from the n-dimension vector y by linear minimum variance estimation is still possible since it exploits (and requires) some *a priori* knowledge on u, i.e., its *a priori* second-order statistic description on Ω_v. Note that, by considering Ω_v to be finer and finer, this prior will be closer and closer to that of a time-continuous input and, accordingly, vector u will determine a piecewise profile closer and closer to a time-continuous function. In other words, this procedure exploits the *a priori* knowledge about the input and its continuity properties directly in its natural continuous-time domain.

C. A Stochastic Model Describing ISR during IVGTT

A glucose stimulus is known to cause an initial spike in ISR followed by a much more regular pattern. In this paragraph we propose a stochastic model which formalizes this *a priori* knowledge.

The model considers the input on a grid divided, for sake of reasoning, in two portions of length N_1 and N_2, i.e., $\Omega_v = \Omega_{v1}, \Omega_{v2}\}$. In the first portion the model must account for the spiky insulin release where the time-course of ISR changes very quickly. Ω_{v1} ends at min 12 or min 15, i.e., $T_{N1} = 12$ or 15, depending on the subject. Each sample of ISR on Ω_{v1} is completely unpredictable given the previous one. One may think that, in this portion, the *a priori* covariance matrix of the process is infinite: only the *a posteriori* information is used to reconstruct the signal in this portion, so Ω_{v1} overlaps the sampling grid Ω_s in its first N_1 elements. The estimation of u_k for $k \leq N_1$ is completely determined by the first N_1 observations (number of unknowns = number of independent equations). The second portion of the virtual grid can be much finer than the corresponding sampling grid. Here we considered a 1 min evenly-spaced grid, i.e., $\Omega_{v2} = \{T_{N1} + 1, T_{N1} + 2, \cdots, 240\}$. In this portion the unknown input is known to be regular and we assumed it to be described by a random-walk model

$$u_k = u_{k-1} + w_k \quad k > N_1 \qquad (27)$$

where w_k is a white noise process with zero mean and variance λ^2. The random-walk model driven by a zero-mean white noise is commonly used to describe smooth signals, e.g., [10], since the difference between two consecutive samples is a random variable with finite variance λ^2. Note that the lower is λ^2, the more regular is process (27).

Broadly speaking, the covariance matrix Σ_u can be thought as

$$\Sigma_u = \lambda^2 \begin{bmatrix} \infty & 0 \\ 0 & (P^T P)^{-1} \end{bmatrix} = \lambda^2 R. \qquad (28)$$

The block structure of the covariance matrix Σ_u denotes the uncorrelation between the two regions where *a priori* knowledge about ISR is available and where it is not. The lower-right block $\lambda^2(P^T P)^{-1}$ corresponds to the covariance matrix of the portion of vector u described by the random-walk model. In particular, matrix P is a N_2-dimension Toeplitz square matrix whose first column is $[1, -1, 0, \cdots, 0]^T$; see e.g., [10].

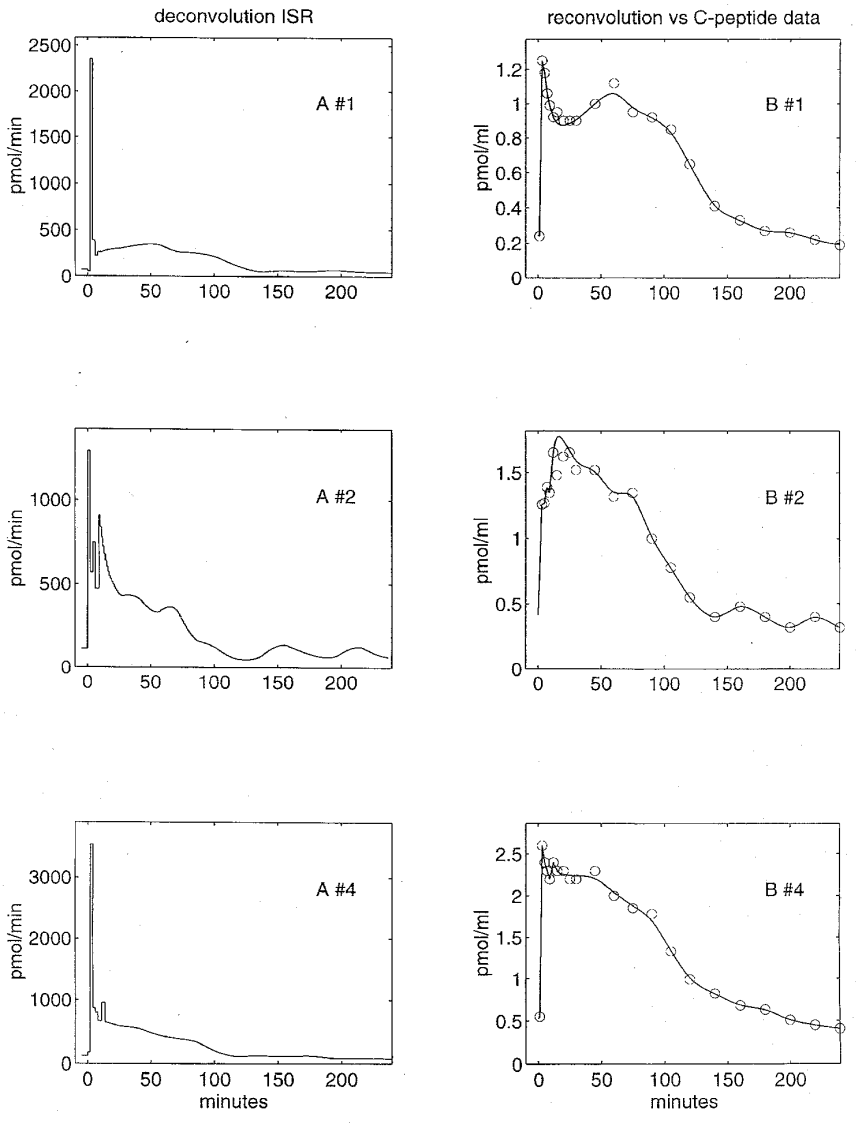
In theory, Ω_{v2} could be chosen even much finer. Note that when a finer and finer Ω_{v2} is selected, the random-walk becomes closer and closer to the integral of a time-continuous white noise process and the prior (27), made on the input first discrete-time differences, becomes closer and closer to a prior on the input first continuous-time derivatives. Finally, it is worth noting that other models, e.g., an integrated random-walk model, could have been used to describe the second-phase ISR as well.

D. The Choice of the Regularization Parameter

Consider $\Sigma_v = \sigma^2 B$ where σ is equal to CV (assumed constant in each data set) and B is a diagonal matrix whose entries are the squared measurements. The covariance matrix of the unknown vector u, i.e., $\Sigma_u = \lambda^2 R$, is known except for the scale factor λ^2.

If the value of σ^2 is assumed to be known, only λ^2 has to be adjusted to obtain the minimum variance estimate: for its choice one has to face problems equivalent to the ones encountered with the regularization parameter (Section V-C). For example, one may resort to Criterion 1, but, since this criterion has an intuitive motivation but not firm statistical grounds, the statistical properties of the estimate mentioned in Section VI-A could not be properly invoked. On the other hand, other available criteria, e.g., generalized cross-validation [24], L-curve [25], have been developed on a somewhat heuristic basis and/or in a deterministic setting as well and thus their use in a stochastic context is questionable too.

Here our *a priori* knowledge about measurement error is also approximate. We only know that CV should range



(a)

Fig. 5. The stochastic approach. (a) *Left panel:* Deconvoluted ISR. *Right panel:* Model predictions against data (subjects #1, #2, and #4).

between 4–6%, i.e., σ should range between 0.04 and 0.06. The individual tuning of the CV is, however, important. In fact, adopting the same CV in all the subjects, say 5%, determines too much smoothing in some individuals, and too little in others (results not shown). In addition, in the impulse response study we have noted that the *a posteriori* CV estimate (Table I) tends to vary, albeit in a small range, among individuals. Therefore, our strategy was to treat the CV 4–6% information as indicative and to simultaneously estimate the input together with the so-called statistical power of both the measurement error (i.e., σ^2) and unknown signal (i.e., λ^2).

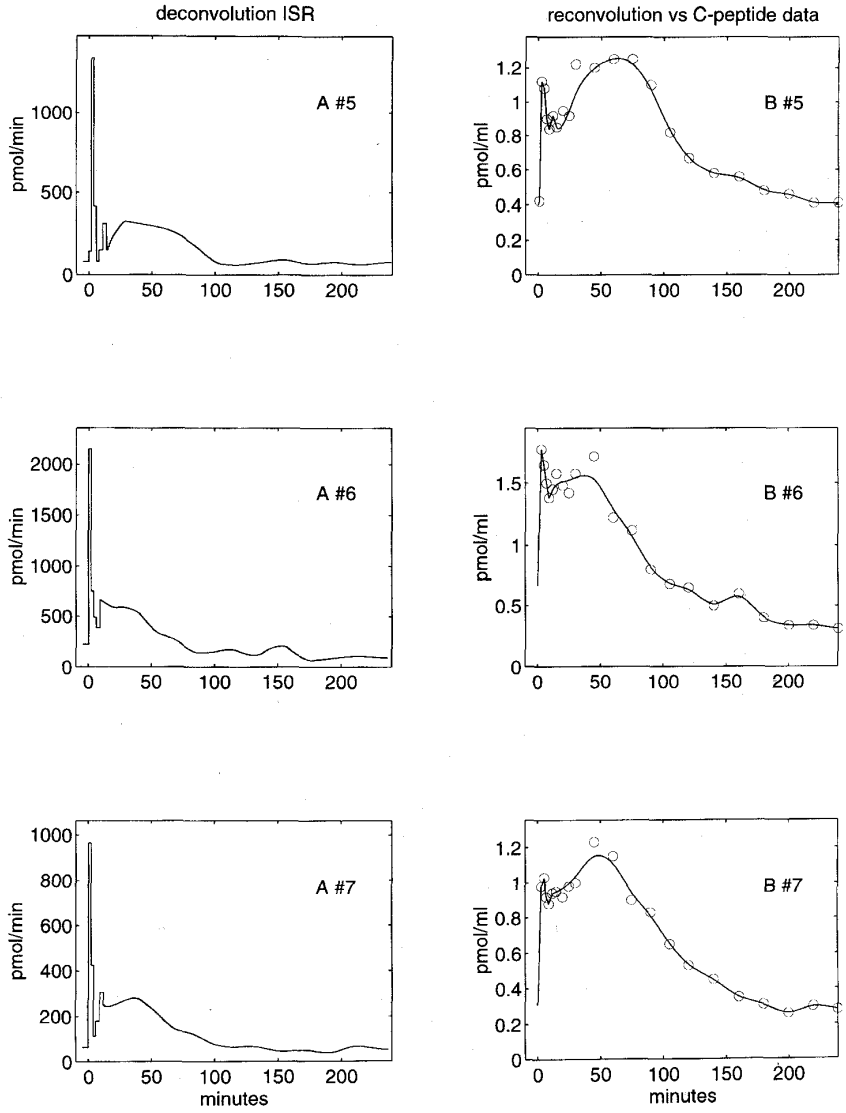
Exploiting the stochastic embedding, statistically-based criteria can be derived for the choice of λ^2 and σ^2. In fact, for the linear minimum variance estimate of problem (21), it holds

$$E\left[\mathrm{WESS}\left(\frac{\sigma^2}{\lambda^2}\right)\right] = \lambda^2 q\left(\frac{\sigma^2}{\lambda^2}\right) \tag{29}$$

$$E\left[\mathrm{WRSS}\left(\frac{\sigma^2}{\lambda^2}\right)\right] = \sigma^2\left[n - q\left(\frac{\sigma^2}{\lambda^2}\right)\right] \tag{30}$$

where $\mathrm{WRSS} = (y-G\hat{u})^T B^{-1}(y-G\hat{u})$ denotes the weighted residual sum of squares, $\mathrm{WESS} = \hat{u}^T R^{-1}\hat{u}$ denotes the weighted estimate sum of squares, and

$$q(\gamma) = \mathrm{trace}[B^{-1/2}G(G^T B^{-1}G + \gamma R^{-1})^{-1}G^T B^{-1/2}] \tag{31}$$



(b)

Fig. 5 (*Continued.*) (b) As (a) for subjects #5–7.

where $B^{-1/2}$ is such that $B^{-1} = B^{-1/2}B^{-1/2}$. The proofs of (29) and (30) are reported in Appendix A. Given the analogy of (30) and a property of linear regression models [5], where the averaged sum of the squared residuals is a biased estimator of the error variance, with the bias depending on the (integer) number of degrees of freedom of the model, $q(\sigma^2/\lambda^2)$ is named *degree of freedom* associated with the ratio σ^2/λ^2. Here the quantity $q(\sigma^2/\lambda^2)$ is a real number varying from zero to n.

Properties (29) and (30) suggest consistency criteria to choose the parameters λ^2 and/or σ^2 when they are unknown. Precisely, when λ^2 is unknown (σ^2 assumed to be known), one should tune it until WESS equals its statistical expectation (29). Conversely, when σ^2 is unknown (λ^2 assumed to be known), one should tune it until WRSS equals its statistical expectation (30); this case is mainly speculative because λ^2 is usually unknown. When both σ^2 and λ^2 are unknown, one can adopt the following criterion, being $\gamma = \sigma^2/\lambda^2$

Criterion 2: Tune γ until

$$\frac{\mathrm{WRSS}(\gamma)}{n - q(\gamma)} = \gamma \frac{\mathrm{WESS}(\gamma)}{q(\gamma)}. \tag{32}$$

As γ is determined, the estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{\mathrm{WRSS}}{n - q(\gamma)} \tag{33}$$

according to (30).

It is easy to see that, since $q(\gamma) > 0$ for $\gamma > 0$, Criterion 1 is not consistent with (30) and, in a stochastic context, would lead, on the average, to oversmoothing.

E. Results

In each subject we considered both σ^2 and λ^2 unknown and tuned their ratio according to Criterion 2. The estimation algorithm was implemented by exploiting suitably designed numerical tools [14]. In Fig. 5 we present, for all subjects but #3, the deconvoluted profile and the corresponding model predictions together with the data. The deconvoluted profile of subject #3 is presented in the next section together with its confidence interval. In some cases the sample at time 1 min was an outlier and was eliminated from the data set. Ω_{v2} starts at min 12 in subjects #2, #3, #6 and at min 15 in subjects #1, #4, #5, #7. The *a posteriori* estimate of the measurement error CV of the IVGTT data, obtained by (33), are, for subject #1–7: 3.9%, 6, 3.7, 2.7, 7.4, 7.1, and 5.2. They vary in a range consistent with both our *a priori* knowledge and the *a posteriori* estimate of Table I.

Having accounted for *a priori* knowledge through (27) and (28), the initial peak is estimated without systematic oversmoothing. Thanks to the introduction of the virtual grid, the second-phase ISR profile is practically time-continuous and more physiologically realistic. Finally, Criterion 2 seems to indicate in each subject a reasonable compromise between adherence to data and regularization.

Note that "linear regularization" was sufficient to obtain nonnegative ISR profiles in all subjects. Would had this not happened one should either develop a stochastic model more complex than e.g., (27) and use it in conjunction with nonlinear estimation algorithms, see e.g., [13], or pose the input estimation problem in a deterministic context, e.g., [10], [12], and [52].

It is possible to make a qualitative comparison between the average secretory profile, displayed in Fig. 6, and the one presented in Fig. 1 of [41] obtained from the same data set. Our estimate gives a sharper and higher peak whereas, in the second part, the mean levels are substantially equivalent. Of note is that our average profile is smoother and shows, from min 60, a slow, low in amplitude, oscillatory pattern. However, it is worth remarking that here a three instead of a two-exponential impulse response model was used. The different spectral content of the lower order model is likely to affect the peak amplitude estimation, especially in those subjects where the impulse response is better described by a three exponential model [43].

F. Confidence Limits

It is important to derive the confidence limits of the insulin secretory profiles reconstructed by deconvolution. There are two sources of uncertainty. The first is the measurement error of the CP concentration measured during IVGTT. The second source is related to the impulse response of the CP system, i.e., the parameters of the impulse response are estimated with a certain precision. For sake of reasoning, we name the first source of uncertainty as data noise, the second as parameter uncertainty.

As already pointed out in [10] and [13], in a stochastic context the covariance matrix of the estimation error can be

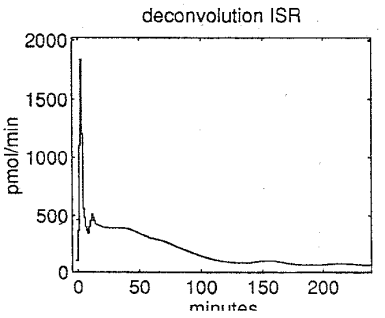


Fig. 6. Average of the estimated ISR.

analytically computed by assuming no parameter uncertainty. It is worth noting that, provided that σ^2 and λ^2 agree with their statistical meaning, the matrix given by (23) accounts for bias error also, see e.g., [27], whereas, in the deterministic context, a statistical characterization of the bias error affecting the estimate cannot be derived [28]. In the Gaussian case, the diagonal elements of matrix (23) allow the computation of the confidence intervals for the entries of $\hat{u}$. In the non-Gaussian case, only standard intervals [19] can be obtained through (23).

However, if one desires a confidence region which also accounts for parameter uncertainty, analytical expressions are not available. Note that, in some cases, e.g., [6], parameter uncertainty can be the major determinant of the precision of the deconvoluted profile. To take into account both errors a natural candidate tool is the Monte-Carlo simulation, see e.g., [18]–[20], which allows us to substitute a considerable computational effort in place of (difficult) theoretical analysis. The strategy is similar to that already used in [6]. In each subject, 300 decay curves were obtained by adding artificial noise (having normal distribution and CV taken from Table I) to the nominal impulse response (parameters taken from Table I) and 300 three-exponential models were identified. Next, 300 IVGTT data sets were obtained by adding artificial noise (having normal distribution and CV as in Section VI-E) to the nominal model predictions (i.e., the reconvolution, as in Fig. 5). We then coupled each simulated IVGTT data set with a set of impulse response parameters and performed deconvolution 300 times. The dispersion of the so-obtained 300 deconvoluted profiles depends on both data noise and parameter uncertainty.

In Fig. 7(a) we show, for subject #3, the (nominal) deconvoluted ISR profile together with its 95% confidence intervals obtained from the 300 runs by a percentile method, i.e., the 95% confidence interval is the interval between the 2.5 and 97.5 percentiles of the Monte-Carlo distribution of the estimated input at each time point [18], [57], [58]. Fig. 7(b) shows the mean $\pm$ SD of the percentage estimation error calculated pointwise in each Monte-Carlo run as: $100 \times$ (nominal ISR-estimated ISR)/(nominal ISR). Finally, Fig. 7(c) shows the model predictions together with their 95% confidence intervals against the data. In the first phase (no regularization) the mean percentage error [solid line, (b)] lies between $\pm 2\%$. In the second phase, the mean percentage error is slightly biased in correspondence of the ISR oscillations (-3% at time 120,

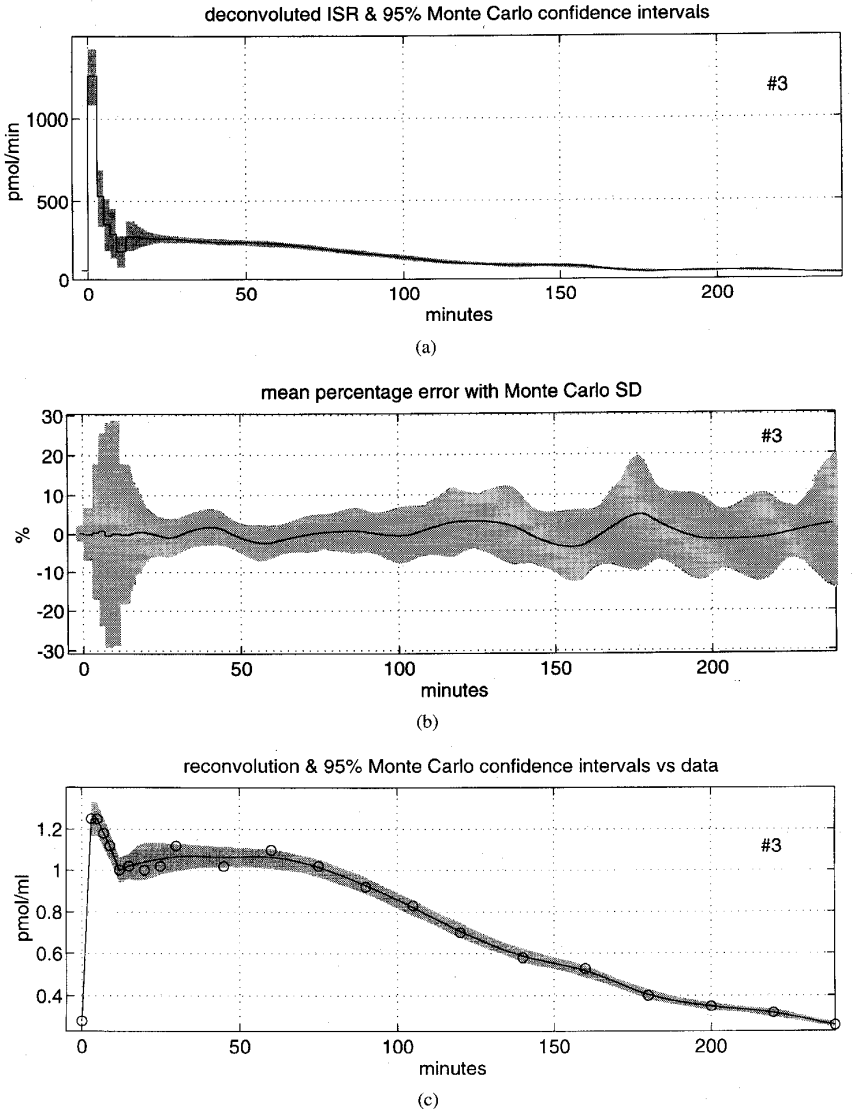


Fig. 7. Confidence intervals in a representative study (subject #3). (a) Deconvoluted ISR together with its 95% confidence intervals. (b) Mean percentage error $\pm$ SD. (c) Model predictions together with their 95% confidence intervals against data.

+3% at time 160, −4% at time 180). Rather expectedly, the largest variability of the percentage estimation error [shaded area, Fig. 7(b)] occurs on Ω_1 (no *a priori* information is specified and, in each run, the estimator must believe to the data). The percentage estimation error SD, initially equal to 2% (reconstruction of basal secretion), raises from 8% (major peak amplitude) up to 30% (interval between times five and 12). Then it has a rapid decrease followed by a somewhat steady 5% value maintained till 70 min. Finally, from 70 min on, there is an increase of the SD up to 20%, also according to the less frequent sampling (less *a posteriori* information is available).

The Monte-Carlo procedure was also repeated to study in each subject the individual effect of the two sources of uncertainty on the precision of the deconvoluted profile. First, the 300 IVGTT data sets were deconvoluted using the nominal impulse response, thus obtaining a dispersion due to data noise only. Then, the nominal model predictions were deconvoluted with the 300 noisy impulse responses, thus obtaining profiles whose dispersion is due to parameter uncertainty only. The sample variance obtained when both errors are considered was seen to be approximately the sum of the two sample variances obtained by considering each source separately. Where the ISR is very smooth, say from min 30 on for subject #3, the contribution of parameter uncertainty is very minor (for subject #3: from 7% of the global variance at 30 min down to 2% at 240 min). Its role in the global uncertainty is, however, important when ISR changes rapidly (for subject #3: 50%

of the global variance at the first-phase peak), also given the lesser precision of the parameters describing the fastest component of the impulse response (see Table I).

Similar results have been obtained for all the other subjects.

VII. Conclusions

Deconvolution allows the indirect measurement of ISR during IVGTT by solving an input estimation problem. The major difficulties to be addressed are the nonuniform/infrequent sampling, the choice of the amount of regularization and the signal nonstationarity.

In this paper, the introduction of the virtual grid, by using an available *a priori* knowledge on the continuity properties of the unknown input in the second-phase ISR, eliminates the staircase approximation and provides secretion rate profiles which are physiologically more realistic. Of course, the use of the virtual grid is helpful to exploit the *a priori* knowledge on the input but it can not be a panacea for possible loss of information caused by infrequent sampling.

The formulation of the problem in a stochastic embedding is not completely new, see e.g., [10] and [13]. Here it has allowed us to derive statistically-based criteria for choosing the amount of regularization. The use of Criterion 2 has indicated an appropriate regularization in each subject and provided a reasonable *a posteriori* estimate for the measurement error CV. Confirmatory results have also come from a test on a simulation problem (Appendix B). A possible additional advantage of the Bayesian embedding consists in providing, assuming an error-free impulse response model, a closed form expression to compute the confidence intervals. Of course, the possibility of invoking the optimal properties of the estimator depends on how reliable the statistical priors adopted to describe the unknown input are. Here, the goodness of (27) and (28) has been assessed *a posteriori* by the qualitative/quantitative results as well as by using the method in a simulation context (Appendix B).

As far as nonstationarity of ISR pattern is concerned, the use of model (28) has lead to an estimate of the initial peak without systematic oversmoothing. Signal nonstationarity could have also been handled differently. We have chosen a solution which maintains a sort of statistical flavor, the aim being to take advantage of the stochastic embedding. From a deterministic point of view, the use of (28) is equivalent to apply no regularization on the first portion of the data to get a complete adherence to them and to concentrate the smoothing action only in the second part of ISR (known to be regular). This avoids wide supposedly spurious oscillations to appear in the slow part of the estimated ISR profile, as it happens when all the regularization effort is played at the beginning. Note that initial "null residuals" (Fig. 5) are not inconsistent with the assumption of a constant measurement error CV but only denote that, when there is a lack of *a priori* knowledge, the linear minimum variance estimator can do nothing but to "believe" to the data (*a posteriori* knowledge). On the contrary, with the "deterministic" algorithm (Fig. 4) the inconsistency of the residuals (large in the beginning, null in the second phase) with the assumption of constant CV was due to the systematic oversmoothing of the first-phase ISR. Of course, null residuals entail the risk to explain both noise and data. Indeed, the high sensitivity of the first-phase ISR to data noise is reflected by a larger confidence region, see e.g., Fig. 7.

To conclude, the algorithm presented in this paper was developed to solve a specific case study, i.e., the reconstruction of ISR after a glucose stimulus. However, while the covariance model (28) is problem oriented, the approach machinery, e.g., the new regularization criteria and the virtual grid, are general purpose.

Appendix A

Below we give the proof of (29) and (30). Consider model (17) and assume v and u uncorrelated with covariance matrices given by

$$\begin{aligned}\Sigma_v &= \sigma^2 B \\ &= \sigma^2 B^{1/2} B^{1/2} \qquad (34)\\ \Sigma_u &= \lambda^2 R \\ &= \lambda^2 (F^T F)^{-1} \qquad (35)\end{aligned}$$

where $B^{1/2}$ and F are two full-rank square matrices of dimension n and N, respectively.

Let us begin with the proof of (30). The linear minimum variance estimator of u given y is given by (22); equivalently by means of the matrix inversion lemma one has [2]

$$\hat{u} = RG^T(GRG^T + \gamma B)^{-1} y \qquad (36)$$

where $\gamma = \sigma^2/\lambda^2$.

The residuals vector, $r = y - G\hat{u}$, is then

$$r = [I_n - GRG^T(GRG^T + \gamma B)^{-1}]y \qquad (37)$$

where I_n denotes the n-dimension eye matrix.

Reminding that the covariance matrix of y is

$$\operatorname{cov}(y) = \lambda^2(GRG^T + \gamma B) \qquad (38)$$

by means of the matrix inversion lemma, one finds

$$\operatorname{cov}(r) = \sigma^2[B - G(G^T B^{-1} G + \gamma R^{-1})^{-1} G^T]. \qquad (39)$$

Defining

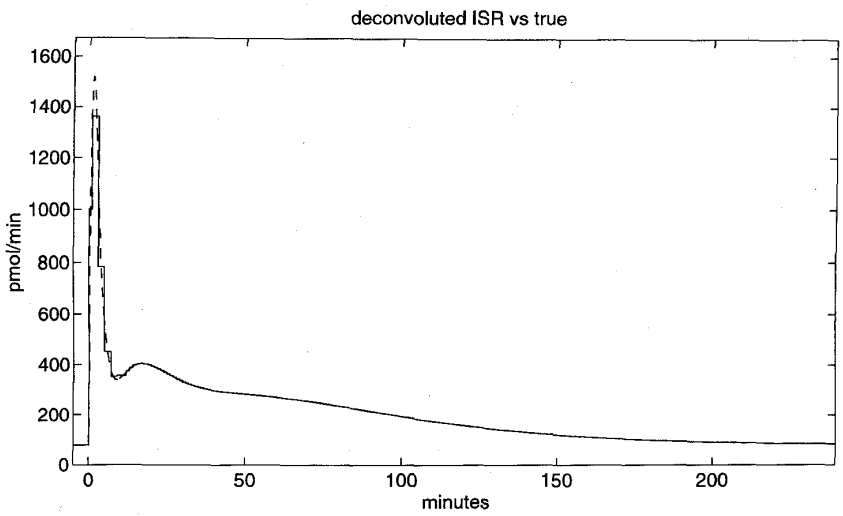
$$r_w = B^{-1/2} r \qquad (40)$$

it holds

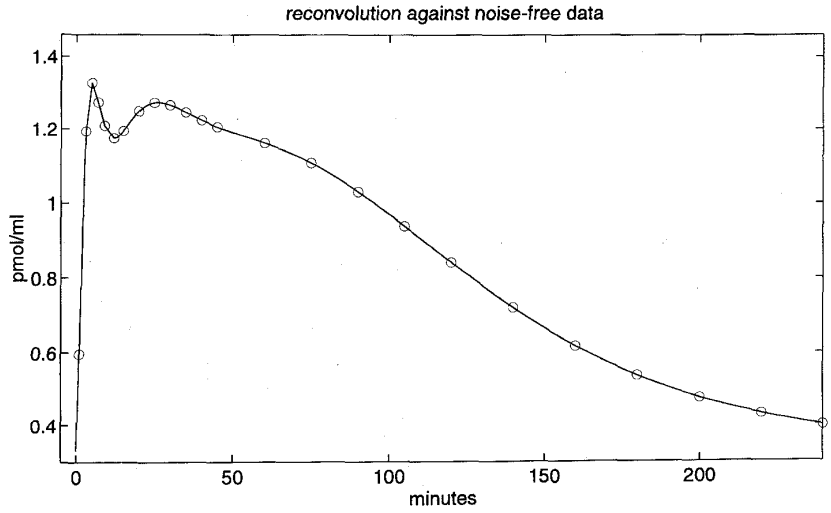
$$\begin{aligned}\operatorname{cov}(r_w) = \sigma^2[I_n - B^{-1/2} G (G^T B^{-1} G \\ + \gamma R^{-1})^{-1} G^T B^{-1/2}]. \qquad (41)\end{aligned}$$

Thereby, we have

$$\begin{aligned}E[\text{WRSS}] &= E[r^T B^{-1} r] \\ &= E[r_w^T r_w] \\ &= \operatorname{trace}[\operatorname{cov}(r_w)] \\ &= \sigma^2[n - q(\gamma)] \qquad (42)\end{aligned}$$



(a)



(b)

Fig. 8. Simulation study. (a) Deconvoluted (continuous line) against true ISR (dashed line). (b) Model predictions against data.

where

$$q(\gamma) = \text{trace}\,[B^{-1/2}G(G^TB^{-1}G + \gamma R^{-1})^{-1}G^TB^{-1/2}] \tag{43}$$

such that (30) is proven.

Now, let us prove (29). From (36), having recalled (38) and (35), it follows

$$\begin{aligned}\text{cov}\,(\hat{u}) &= \lambda^2 RG^T(GRG^T + \gamma B)^{-1}GR \\ &= \lambda^2 F^{-1}F^{-T}G^T(GF^{-1}F^{-T}G^T + \gamma B)^{-1} \\ &\quad \cdot GF^{-1}F^{-T}.\end{aligned} \tag{44}$$

Hence, having defined

$$\hat{u}_w = F\hat{u} \tag{45}$$

it holds

$$\text{cov}\,(\hat{u}_w) = \lambda^2 F^{-T}G^T(GF^{-1}F^{-T}G^T + \gamma B)^{-1}GF^{-1}. \tag{46}$$

In order to easily manipulate (46), let us define

$$H = B^{-1/2}GF^{-1}. \tag{47}$$

H is a $n \times N$ matrix. From (46) and (47) we thus have

$$\begin{aligned}\text{cov}\,(\hat{u}_w) &= \lambda^2 F^{-T}G^TB^{-1/2}(B^{-1/2}GF^{-1} \\ &\quad \cdot F^{-T}G^TB^{-1/2} + \gamma I_n)^{-1}B^{-1/2}GF^{-1} \\ &= \lambda^2 H^T(HH^T + \gamma I_n)^{-1}H.\end{aligned} \tag{48}$$

Let U and V be unitary matrices, of dimension n and N, respectively, such that

$$U^THV = D. \tag{49}$$

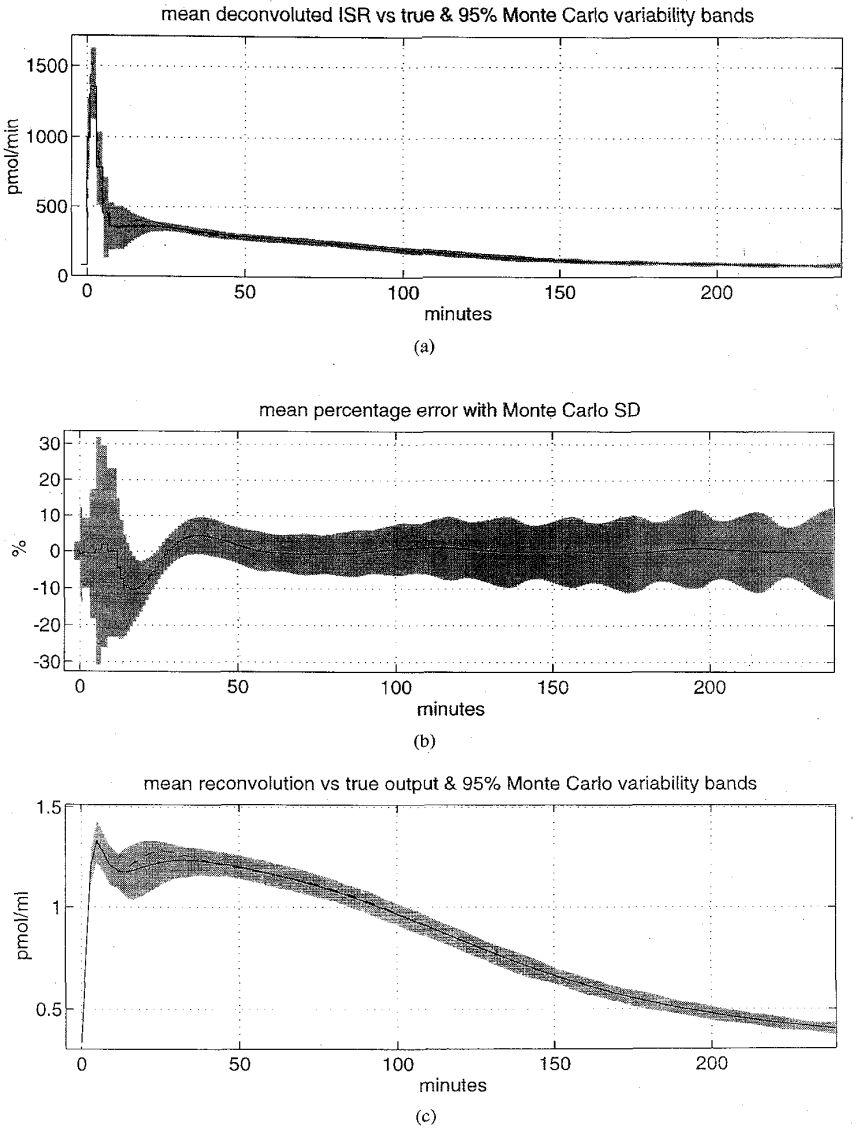


Fig. 9. Simulation study. (a) True ISR (dashed line), mean deconvoluted ISR (continuous line), and 95% variability bands (shaded area). (b) Mean percentage error ± SD. (c) True concentration profile (dashed line), mean reconvolution (continuous line), and 95% variability bands (shaded area).

D being a diagonal $n \times N$ rectangular matrix whose n diagonal elements are denoted with d_i, $i = 1, 2, \cdots, n$. It follows that

$$\begin{aligned}\text{trace}\,[H^T(HH^T+\gamma I_n)^{-1}H] &= \text{trace}\,[VD^TU^T(UDV^T \\ &\quad \cdot VD^TU^T+\gamma I_n)^{-1}UDV^T] \\ &= \text{trace}\,[D^T(DD^T+\gamma I_n)^{-1}D] \\ &= \sum_{i=1}^{n}\frac{d_i^2}{d_i^2+\gamma}. \end{aligned} \tag{50}$$

On the other hand, from (43), (47), and (49), it also holds

$$\begin{aligned} q(\gamma) &= \text{trace}\,[B^{-1/2}G(G^TB^{-1}G+\gamma R^{-1})^{-1}G^TB^{-1/2}] \\ &= \text{trace}\,[B^{-1/2}GF^{-1}(F^{-T}G^TB^{-1/2}B^{-1/2}GF^{-1} \\ &\quad + \gamma I_N)^{-1}F^{-T}G^TB^{-1/2}] \\ &= \text{trace}\,[H(H^TH+\gamma I_N)^{-1}H^T] \\ &= \text{trace}\,[U^TDV(V^TD^TUU^TDV+\gamma I_N)^{-1}V^TD^TU] \\ &= \text{trace}\,[D(D^TD+\gamma I_N)^{-1}D^T] \\ &= \sum_{i=1}^{n}\frac{d_i^2}{d_i^2+\gamma}. \end{aligned} \tag{51}$$

From (48), by comparing (50) and (51), we have finally

$$\begin{aligned} E[\text{WESS}] &= E[\hat{u}^TR^{-1}\hat{u}] \\ &= E[\hat{u}_w^T\hat{u}_w] \\ &= \text{trace}\,[\text{cov}\,(\hat{u}_w)] \\ &= \lambda^2 q(\gamma). \end{aligned} \tag{52}$$

APPENDIX B

In this appendix, we show the performance of the method in a simulation study. The function

$$\mathrm{ISR}(t) = \begin{cases} C & t < 0 \\ C + B_1 e^{-b_1 t} + B_2 e^{-b_2 t} & \\ \quad + B_3 t^{p_1} e^{-b_3 t} + B_4 t^{p_2} e^{-b_4 t} & t \geq 0 \end{cases} \tag{53}$$

with $C = 80$, $B_1 = -B_2 = 8000$, $B_3 = 1.4$, $B_4 = 1.2$, $p_1 = 3$, $p_2 = 1.75$, $b_1 = 1.8^{-1}$, $b_2 = 1.1^{-1}$, $b_3 = p_1/15$, and $b_4 = p_2/50$, simulates an ISR profile during an IVGTT. The function

$$g(t) = \sum_{i=1}^{3} A_i e^{-\alpha_i t} \tag{54}$$

with $A_1 = 1.794 \times 10^{-4}$, $A_2 = 8.801 \times 10^{-5}$, $A_3 = 4.954 \times 10^{-5}$ $a_1 = 0.313$, $a_2 = 0.073$, $a_3 = 0.021$, represents the impulse response of the CP system. The function $c(t) = g(t) \otimes \mathrm{ISR}(t)$, where "$\otimes$" denotes convolution, thus gives the CP plasma concentration time-course during the simulated IVGTT.

Parameters of (53) and (54) have been chosen by the following rationale. The input profile, shown in Fig. 8, was chosen to bear a qualitative resemblance with the average estimated ISR profile reported in [41]. The impulse response parameters were obtained by fitting a three exponential sum to the mean of the decay curves.

First, we considered the ideal situation, where the true system impulse response $g(t)$ is known and the noise-free samples of $c(t)$ with the schedule of Section III are available. In such a case no regularization is needed to solve the inverse problem. The unknown input was described by (27) and (28) on the grid Ω_v of Section VI-C (Ω_{v2} starts at min 13). Fig. 8(a) shows the reconstruction of the ISR profile (continuous line) together with the true ISR (dashed line). Fig. 8(b) shows the model predictions, i.e., reconvolution, against the simulated noise-free data. In the first phase, given the finite sampling rate and the lack of *a priori* information, the approach can not recover pointwise the true ISR and the values of the staircase profile should be interpreted as mean of the input in each sampling interval. In the second phase, *a priori* information is available and the true time-continuous profile is perfectly reconstructed.

To evaluate the performance of the method in a more realistic situation we considered 300 simulations of the two-stage experiment of Section III by generating noisy samples of the response to a CP bolus and noisy measurements of the CP concentration during the IVGTT. The artificial measurement error is additive, Gaussian and uncorrelated with a constant CV = 5%. For each of the 300 simulations, we first fitted a three exponential model; then this model was used in the deconvolution of the IVGTT data having assumed both σ^2 and λ^2 unknown. Fig. 9(a) shows the true ISR profile (dashed line), the mean of the 300 deconvoluted ISR profiles (continuous line) together with the 95% variability bands (shaded area), i.e., the interval between the 2.5 and 97.5 percentiles of the Monte-Carlo distribution of the estimated input at each time point. Fig. 9(b) shows the mean and SD of the percentage errors, calculated pointwise in each run as in Section VI-F. Fig. 9(c) displays the true concentration profile (dashed line), the mean of the 300 reconvolution profiles (continuous line) together with the 95% variability bands (shaded area). The average estimated ISR is unbiased except near the second peak at min 15 (where the bias is around 10%) where the input must be reconstructed from few samples (at 12, 15, 20, and 25 min) of a small amplitude oscillation corrupted by a relatively large measurement error, i.e., around min 15 there is a low signal-to-noise ratio (SNR). Rather expectedly, the largest variability of the percentage estimation error [shaded area, (b)] occurs in Ω_{v1}. The percentage estimation error SD, initially equal to 2.7% (reconstruction of basal secretion), raises from 10% (major peak amplitude) up to 38% and then it has a rapid decrease followed by a somewhat steady value (around 7%) maintained from 20–70 min. Finally, from 70–240 min, SD raises up to 13%, according to the less frequent sampling.

Interestingly, the mean of the 300 *a posteriori* estimated CV's, obtained in each single run by Criterion 2, was $5.1 \pm 1.5(\%)$, suggesting that the used criterion provides suitable regularization.

The Monte-Carlo procedure was also repeated two more times. First, the 300 IVGTT noisy data sets were deconvoluted using the true impulse response. Then, the noise-free IVGTT samples were deconvoluted with the 300 perturbed impulse responses. Again, the sample variance obtained when both the errors are simultaneously considered was seen to be approximately the sum of the two sampled variances obtained considering each source separately. Where ISR is very smooth, say from min 40 on, the contribution of the impulse response parameter uncertainty is very minor (from 5% of the global variance to 2% at 240 min). Its role in the global uncertainty is, however, important around the initial ISR peak (up to 42% of the global variance).

ACKNOWLEDGMENT

The authors thank Dr. K. Polonsky (Department of Medicine, University of Chicago, IL) for having provided the data published in [41]. The authors are grateful to Dr. G. De Nicolao for his helpful suggestions.

REFERENCES

[1] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Control*, vol. AC-19, pp. 716–723, 1974.
[2] J. V. Beck and K. J. Arnold, *Parameter Estimation in Engineering and Science*. New York: Wiley, 1977.
[3] R. N. Bergman, D. T. Finegood, and M. Ader, "Assessment of insulin sensitivity *in vivo*," *Endocrinol. Rev.*, vol. 6, pp. 45–86, 1985.
[4] M. Bertero, "Linear inverse and ill-posed problems," *Advances in Electronics and Electron Phys.*, vol. 75, pp. 1–120, 1989.
[5] E. Carson, C. Cobelli, and L. Finkelstein, *The Mathematical Modeling of Metabolic and Endocrine Systems*. New York: Wiley, 1983.
[6] A. Caumo and C. Cobelli, "Hepatic glucose production during IVGTT: Estimation by deconvolution with a new minimal model," *Am. J. Physiol.*, vol. 264, pp. 829–841, 1993.
[7] M. K. Charter and S. F. Gull, "Maximum entropy and its application to the calculation of drug absorption rates," *J. Pharmacokinet. Biopharm.*, vol. 15, pp. 645–655, 1987.
[8] ——, "Maximum entropy and drug absorption," *J. Pharmacokinet. Biopharm.*, vol. 19, pp. 497–520, 1991.

[9] C. Cobelli, A. Mari, S. Del Prato, S. De Kreutzenberg, R. Nosadini, and J. Jensen, "Reconstructing the rate of appearance of subcutaneous insulin by deconvolution," *Am. J. Physiol.,* vol. 252, pp. E549–E556, 1987.

[10] D. Commenges, "The deconvolution problem: Fast algorithms including the preconditioned conjugate-gradient to compute a MAP estimator," *IEEE Trans. Automat. Control,* vol. AC-29, no. 3, pp. 229–243, 1984.

[11] D. J. Cutler, "Numerical deconvolution by least squares: Use of prescribed input functions," *J. Pharmacokin. Biopharm.,* vol. 6, no. 3, pp. 227–241, 1978.

[12] G. De Nicolao and M. Rocchetti, "Stable and efficient techniques for the deconvolution of hormone time-series," in *Computers in Endocrinology: Recent Advances,* G. Forti, V. Guardabasso, and D. Rodbard, Eds. New York: Raven, 1990, pp. 83–91.

[13] G. De Nicolao and D. Liberati, "Linear and nonlinear techniques for the deconvolution of hormone time-series," *IEEE Trans. Biomed. Eng.,* vol. 40, pp. 440–455, 1993.

[14] G. De Nicolao, "Numerical deconvolution of physiological systems: A stochastic approach," in *Proceedings of the IFAC Symposium on Modeling and Control in Biomedical Systems,* Galveston, TX, B. W. Patterson, Ed. Madison, WI: Omnipress, 1994, pp. 242–253.

[15] G. De Nicolao, D. Liberati, and A. Sartorio, "Deconvolution of infrequently sampled data for the estimation of growth hormone secretion," *IEEE Trans. Biomed. Eng.,* vol. 42, no. 7, 1995.

[16] D. Donoho, I. M. Johnstone, J. Hoch, and A. Stern, "Maximum entropy and the nearly black object," *J. Roy. Stat. Soc. Ser. B,* vol. 54, pp. 41–81, 1992.

[17] R. P. Eaton, R. C. Allen, D. S. Schade, K. M. Erickson, and J. Standefer, "Prehepatic insulin production in man: Kinetic analysis using peripheral connecting peptide behavior," *J. Clin. Endocrinol. Metab.,* vol. 51, pp. 520–528, 1980.

[18] B. Efron, *The Jacknife, the Bootstrap and Other Resampling Plans.* SIAM CBMS-NSF: Philadelphia, PA, 1982, Monograph no. 38.

[19] B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Stat. Sci.,* vol. 1, pp. 54–77, 1986.

[20] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap.* New York: Chapman & Hall, 1993.

[21] M. P. Ekstroem, "A spectral characterization of the ill-conditioning in numerical deconvolution," *IEEE Trans. Audio Electroacoust.,* vol. AU-21, pp. 344–348, 1973.

[22] O. K. Faber, C. Hagen, C. Binder, J. Markussen, V. K. Nahitani, P. M. Blix, H. Kuzuya, D. L. Horwitz, A. H. Rubenstein, and N. Rossing, "Kinetics of human connecting peptide in normal and diabetic subjects," *J. Clin. Invest.,* vol. 62, pp. 197–202, 1979.

[23] W. Gillespie and P. Veng-Pedersen, "A polyexponential deconvolution method. Evaluation of the gastrointestinal bioavailability and mean *in vivo* dissolution time of some ibuprofen dosage forms," *J. Pharmacokin. Biopharm.,* vol. 13, pp. 289–307, 1985.

[24] G. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics,* vol. 21, pp. 215–224, 1979.

[25] P. C. Hansen, "Numerical tools for analysis and solutions of Fredholm integral equations of the first kind," *Inverse Problems,* vol. 8, pp. 849–872, 1992.

[26] P. Hall and D. M. Titterington, "Common structure of techniques for choosing smoothing parameters in regression problems," *J. Roy. Stat. Soc. Ser. B,* vol. 49, pp. 184–198, 1987.

[27] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models.* London, UK: Chapman & Hall, 1990.

[28] B. R. Hunt, "Biased estimation for nonparametric identification of linear systems," *Math. Biosci.* vol. 10, pp. 215–237, 1971.

[29] ——, "A theorem on the difficulty of numerical deconvolution," *IEEE Trans. Audio Electroacoust.,* vol. AU-20, pp. 94–95, 1972.

[30] E. M. Landaw and J. J. DiStefano, III, "Multiexponential, multicompartmental, and noncompartmental modeling—II: Data analysis and statistical considerations," *Am. J. Physiol.,* vol. 246, pp. R665–R677, 1984.

[31] J. Licinio-Paixao, K. S. Polonsky, B. D. Given, W. Pugh, D. Ostrega, B. H. Frank, and A. H. Rubenstein, "Ingestion of a mixed meal does not affect the metabolic clearance of biosynthetic human C-peptide," *J. Clin. Endocrinol. Metab.,* vol. 63, pp. 401–406, 1986.

[32] V. A. Morozov, "On the solution of functional equations by the method of regularization," *Soviet Math. Dokl.,* vol. 7, pp. 414–416, 1966.

[33] D. Nychka, "Bayesian confidence intervals for smoothing splines," *J. Am. Stat. Assoc.,* vol. 83, pp. 1134–1143, 1988.

[34] F. O'Sullivan, "A statistical perspective on ill-posed inverse problems," *Stat. Sci.,* vol. 1, pp. 502–527, 1986.

[35] D. L. Phillips, "A technique for the numerical solution of certain integral equations of the first kind," *J. Assoc. Comput. Mach.,* vol. 9, pp. 97–101, 1962.

[36] A. Pilo, E. Ferrannini, and R. Navalesi, "Measurement of glucose-induced delivery rate in man by deconvolution analysis," *Am. J. Physiol.,* vol. 233, pp. E500–E508, 1977.

[37] K. S. Polonsky and A. H. Rubenstein, "C-peptide as a measure of the secretion and hepatic extraction of insulin: Pitfalls and limitations," *Diabetes,* vol. 33, pp. 486–494, 1984.

[38] K. S. Polonsky, J. Licinio-Paixao, B. D. Given, W. Pugh, P. Rue, J. Galloway, T. Karrison, and B. Frank, "Use of biosynthetic human C-peptide in the measurement of insulin secretion rates in normal volunteers and type I diabetic patients," *J. Clin. Invest.,* vol. 77, pp. 98–105, 1986.

[39] J. A. Rice, "Choice of the smoothing parameter in deconvolution problems," *Contemp. Math.,* vol. 59, pp. 137–151, 1986.

[40] A. H. Rubenstein, L. A. Pottenger, M. Mako, G. S. Getz, and D. F. Steiner, "The metabolism of insulin and proinsulin by the liver," *J. Clin. Invest.,* vol. 51, pp. 912–920, 1972.

[41] E. T. Shapiro, H. Tillil, A. H. Rubenstein, and K. S. Polonsky, "Peripheral insulin parallels changes in insulin secretion more closely than C-peptide after bolus intravenous glucose administration," *J. Clin. Endocrinol. Metab.,* vol. 67, pp. 1094–1099, 1988.

[42] G. Schwarz, "Estimating the dimension of a model," *Ann. Stat.,* vol. 6, pp. 461–464, 1978.

[43] G. Sparacino and C. Cobelli, "Impulse response model order affects reconstruction of insulin secretion rate by deconvolution. Influence of secretion rate spectral content, sampling frequency and impulse response input design," in *Proceedings of the 3rd European Conference on Engineering and Medicine,* Florence, Italy, A. Pedotti and P. Rabischong, Eds. Milano, Italy: Ed. Pro Juventute Don Gnocchi, 1995, p. 203.

[44] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet. Math. Dokl.,* vol. 4, pp. 1035–1038, 1963.

[45] G. Toffolo, F. De Grandi, and C. Cobelli, "Estimation of beta cell sensitivity from IVGTT C-peptide data. Knowledge of the kinetics avoids errors in modeling the secretion," *Diabetes,* vol. 44, pp. 845–854, 1995.

[46] R. C. Turner, J. A. Grayburn, G. B. Newman, and J. D. N. Nabarro, "Measurement of the insulin delivery rate in man," *J. Clin. Endocrinol. Metab.,* vol. 33, pp. 279–286, 1971.

[47] S. Twomey, "The application of numerical filtering to the solution of integral equations encountered in the indirect sensing measurements," *J. Franklin Inst.,* vol. 279, pp. 95–109, 1965.

[48] S. Vajda, K. R. Godfrey, and P. Valko, "Numerical deconvolution using system identification methods," *J. Pharmacokinet. Biopharm.,* vol. 16, pp. 85–107, 1988.

[49] E. Van Cauter, F. Mestrez, J. Sturis, and K. S. Polonsky, "Estimation of insulin secretion rates from C-peptide levels," *Diabetes,* vol. 41, pp. 368–377, 1992.

[50] J. D. Veldhuis, M. L. Carlson, and M. L. Johnson, "The pituitary gland secretes in bursts: Appraising the nature of glandular secretory impulses by simultaneous multiple-parameter deconvolution of plasma hormone concentration," in *Proc. Natl. Acad. Sci. USA,* 1987, vol. 4, pp. 7686–7690.

[51] P. Veng-Pedersen, "An algorithm and computer program for deconvolution in linear pharmacokinetics," *J. Pharmacokinet. Biopharm.,* vol. 8, pp. 463–481, 1980.

[52] D. Verotta, "An inequality-constrained least-squares deconvolution method," *J. Pharmacokinet. Biopharm.,* vol. 17, pp. 269–289, 1989.

[53] ——, "Comments on two recent deconvolution methods," *J. Pharmacokinet. Biopharm.,* vol. 18, pp. 483–499, 1990.

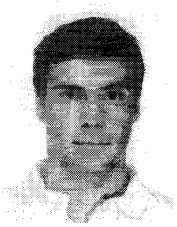
[54] ——, "Estimation and model selection in constrained deconvolution," *Ann. Biomed. Eng.,* vol. 21, pp. 605–620, 1993.

[55] ——, "Two constrained deconvolution methods using spline functions," *J. Pharmacokinet. Biopharm.,* vol. 21, pp. 609–636, 1993.

[56] G. Wahba, "Bayesian "confidence intervals" for the cross-validated smoothing spline," *J. Roy. Stat. Soc. Ser. B,* vol. 45, pp. 133–150, 1983.

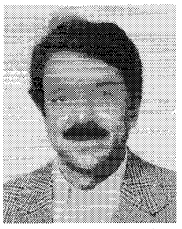
[57] G. Wahba, *Splines models for observational data, SIAM CBMS-NFS,* Reg. Conf. Ser., Philadelphia, PA, 1990.

[58] Y. Wang and G. Wahba, "Bootstrap Confidence Intervals for Smoothing Splines and Their Comparison to Bayesian 'Confidence Intervals,'" Department of Statistics, University of Wisconsin, Madison, WI, Tech. Rep. no. 913, 1994.



Giovanni Sparacino was born in Pordenone, Italy, on November 11, 1967, where he received the high-school diploma in July 1986 from Liceo Scientifico Grigoletti. On May 19, 1992, he received the Doctoral degree (Laurea) in electronic engineering with academic laude from the University of Padova, Padova, Italy. In November, 1992, he had a study grant to attend a three-year Ph.D. program (Dottorato di Ricerca) in biomedical engineering at the University of Padova.

His scientific interests include deconvolution techniques and modeling, simulation and control of metabolic and endocrine systems.



Claudio Cobelli (S'67–M'92) was born in Bressanone (Bolzano), Italy, on February 21, 1946. He received the Doctoral degree (Laurea) in electronic engineering in 1970 from the University of Padova, Padova, Italy.

From 1970–1980, he was a Research Fellow of the Institute of System Science and Biomedical Engineering, National Research Council (LADSEB-CNR), Padova, Italy. From 1973–1975, he was Associate Professor of Biological Systems at the Department of Electronics, University of Florence. From 1975–1981, he was Associate Professor of Biomedical Engineering at the Departments of Electronics, University of Padova. In 1981, he become Full Professor of Biomedical Engineering at the Department of Electronics and Informatics, University of Padova. His main research activity is in the field of modeling and identification of physiological systems, especially endocrine-metabolic and pharmacokinetic systems.

He is co-editor (with R. N. Bergman) of *Carbohydrate Metabolism: Quantitative Physiology and Mathematical Modeling* (Chichester, U.K.: Wiley, 1981), and (with L. Mariani) of *Modeling and Control of Biomedical Systems* (Oxford, U.K.: Pergamon Press, 1989). He is co-author (with E. R. Carson and L. Finkelstein) of *The Mathematical Modeling of Metabolic and Endocrine Systems* (New York: Wiley, 1983). He is on the Editorial Board of *Mathematical Biosciences, American Journal of Physiology, Diabetes, and Nutrition & Metabolism* and Associate Editor of *Diabetologia Control Engineering Practice*. Dr. Cobelli is chairman (1993–1996) of IFAC TC on Modeling and Control of Biomedical Systems. He is member of the International Federation for Medical and Biological Engineering, the Biomedical Engineering Society, the Society for Mathematical Biology, the American Diabetes Association, and the European Association for the Study of Diabetes.