IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 14 September 2022, accepted 27 September 2022, date of publication 5 October 2022, date of current version 12 October 2022.

*Digital Object Identifier 10.1109/ACCESS.2022.3212435*

RESEARCH ARTICLE

# Correlation Method for Identification of a Nonparametric Model of Type 1 Diabetes

**MARTIN DODEK, EVA MIKLOVIČOVÁ, AND MARIÁN TÁRNÍK**
Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, 81219 Bratislava, Slovakia
Corresponding author: Martin Dodek (martin.dodek@stuba.sk)

This work was supported by the Modeling and Control of Biosystems through the Ministry of Education, Science, Development and Sport of the Slovak under Grant VEGA 1/0049/20.

**ABSTRACT** This work describes a novel nonparametric identification method for estimating impulse responses of the general two-input single-output linear system with its target application to the individualization of an empirical model of type 1 diabetes. The proposed algorithm is based on correlation functions and the derived generalization of the Wiener-Hopf equation for systems with two inputs, while taking the stochastic properties of the output measurements into account. Ultimately, this approach to solving the deconvolution problem can be seen as an alternative to widely used prediction error methods. To estimate the impulse response coefficients, the generalized least squares method was used in order to reflect nonuniform variances and nonzero covariances of the stochastic estimate of the cross-correlation functions, hence yielding the minimum variance estimator. Estimate regularization strategies were also involved, while three different types of penalties were applied. The combination of smoothing, stability, and causality regularization was proposed to improve the general validity of the estimate and also to lower its variance. The findings of this identification method are meant to be applied within an eventual predictive control synthesis for the artificial pancreas, so a procedure for transforming the nonparametric model into the transfer function-based parametric model was also described. A discussion on the results of a comprehensive simulation-based experiment concludes the paper.

**INDEX TERMS** Correlation function, generalized least squares method, minimum variance estimate, multiple-input single-output systems, nonparametric model, regularization, system identification, type 1 diabetes mellitus, Wiener-Hopf equation.

## I. INTRODUCTION

Diabetes mellitus is a chronic metabolic disorder characterized by the persistent state of elevated blood glucose concentration, also called hyperglycemia, causing various serious health complications for the patient. In recent decades, this problem has attracted the growing attention of scientists, even among control engineers. In this study, we will focus on type 1 diabetes, which is characterized by absolute insulin deficiency and is therefore considered the insulin-dependent form of this disorder. The main motivation to address the problem of identifying empirical models of diabetes is to improve the accuracy of the prediction of hyper and hypoglycemia states, as these are the major risks associated with diabetes and its treatment.

The individualized model is also essential for correct state estimation [1], [2] and eventual synthesis of a high-performance controller of glycemia [3], [4] that ultimately heads towards implementation of the so-called artificial pancreas [5], [6]. In this field there is still a space for substantial improvement of control performance by means of supplying as valid and accurate model as possible. This model personalization should ideally be performed using only continuous glucose monitoring measurements and diabetic diary logs, i.e., passively acquired data without the need for controlled clinical experiments.

Traditional identification approaches based on minimizing the single step-ahead prediction error that are typically used

The associate editor coordinating the review of this manuscript and approving it for publication was Mauro Gaggero.

in combination with simple autoregressive models are characterized by insufficient performance in prediction-oriented problems and even poorer physiology compliance and general validity. Furthermore, the model structure with a single autoregressive component cannot cover the significant differences between the dynamics of insulin administration and carbohydrate intake. However, there also exist identification approaches and more appropriate model structures that address the aforementioned issues, but these usually lack the analytical closed-form solution of the estimation problem coming at the expense of increased computational complexity or the need for involving the iterative optimization methods. To address these challenges, this paper proposes an identification method that yields a model with fully independent dynamics for both inputs, capable of long-term predicting glycemia, while featuring the parameter estimate in closed form obtained by evaluating the relatively cheap generalized least squares method.

The most significant original contributions presented in this paper can be listed as follows:

- Derivation of the generalized form of the Wiener-Hopf equation for the system with two inputs
- Formulation of the equivalent linear regression system
- Analysis of the statistical properties of the correlation function estimate
- Application of the generalized least squares method to obtain the parameter estimate with minimal variance
- Algorithm for approximating the impulse responses by the transfer function model

### A. STATE OF THE ART

So far, the problem of the so-called patient tailoring or individualizing of mathematical models of glycemia dynamics in subjects with type 1 diabetes has been the subject of extensive research and various in-silico and in-vivo studies, as will be briefly reviewed in this section.

Early findings on basic parametric empirical models such as the autoregressive exogenous (abbr. ARX), autoregressive moving average exogenous (abbr. ARMAX), and the Box-Jenkins (abbr. BJ) models identified using state-of-the-art algorithms minimizing the prediction error [7], were summarized within a comparative study [8] and in a comprehensive overview [9]. As stated in the latter paper, the parameters of the ARMAX and the BJ model had to be estimated from iterative prediction error methods, which are susceptible to convergence to a local minimum due to non-convexity of the corresponding optimization problem. Moreover, most studies have neglected the fact that the two-input ARX model is not structurally compliant with the basic physiology of diabetes due to its shared autoregressive dynamics for both inputs.

A similar comparison of the aforementioned stochastic discrete-time models was presented in [10] where the authors relied on online identification using the recursive least squares method. However, as mentioned in [11], the single step-ahead prediction error does not yield ideal performance in prediction and control-oriented problems. For this reason, the authors also experimented with a promising multiple step-ahead prediction error criterion formed as a nonlinear least squares problem requiring iterative solvers to be involved.

Concerning the class of continuous-time models, another research group in [12], [13], and [14] presented an identification method for a multiple-input single-output transfer function structure that comprised integrators. The problem considered there was to find the model parameters that minimize the quadratic prediction error criterion using the iterative Gauss-Newton algorithm suitable for such nonlinear optimization problems, yet it required the gradient and the Hessian to be supplied. Although this approach is interesting, the model with integrators fails to predict glycemia for longer time periods than just postprandial. A similar strategy, yet for an integrator-less model structure, hand in hand with the derivative-free Nelder-Mead optimization method, was reported in [15].

A different higher-order continuous-time model based on the transfer function was proposed in [16], [17], [18], and [19]

The model was tailored to the patient using the procedure based on minimization of the sum of squared residuals between the measured glycemia and the glycemia obtained by the model simulation. However, the residuals were a nonlinear function of the model parameters, so the nonlinear least squares method had to be used.

In the widely acclaimed work [20] a nonlinear model of glucose kinetics was used in conjunction with the Bayesian parameter estimation adopted to adjust the time-varying model parameters. In detail, identification was carried out by reestimating the model parameters at each iteration employing glucose measurements from the so-called ''learning window''.

In paper [21], a nonlinear system identification approach was proposed based on the prediction error method applied to the Hammerstein Box-Jenkins model structure.

Recent findings in our work [22] have led to an identification strategy that can be characterized as a numeric minimization of the model multiple step-ahead prediction error. The compliance of the identified model with the basic physiology was ensured by performing the identification in the constrained parameter space of the model poles, zeros, and gains.

The last decade has witnessed an increased interest in nonparametric identification methods. The application of conservative prediction error methods for linear models was compared to a novel black-box kernel-based nonparametric approach in [23], [24], and [25]

Paper [26] has attracted our attention because of the linear single-input nonparametric model identified using a cross-correlation method with applied filtering of the impulse-response coefficients via projection onto the Laguerre basis.

For comparison, this study proposes a novel nonparametric approach for estimating impulse responses of the linear

empirical model based on correlation functions and the generalized least squares method. The paper has been divided into the following sections: In section II, the structure of the two-input single-output continuous-time nonparametric model is defined using the convolution integrals and the corresponding impulse functions. This is followed by section III where the essential continuous-time integral equations and the generalized form of the Wiener-Hopf equation are derived. Discrete-time equations for the finite-length experiment are presented in section IV, where the covariance matrix of the estimated cross-correlation functions is also derived. The equivalent linear equation system is formed in section V, where the coefficients of impulse responses are estimated using the generalized least squares method with the estimate regularization. The same section examines the rationale for each of the applied regularization strategies. The auxiliary equations for the parametric model approximation from the identified nonparametric model can be found in section VI. The experimental setup and the results of the simulation-based case study aimed at estimating the parameters of an empirical model of type 1 diabetes are discussed in section VII.

## II. MODEL STRUCTURE AND PRELIMINARIES

The fundamental preliminary is the two-input single-output linear nonparametric model with dynamics defined by its impulse functions. In the context of empirical modeling of glycemia dynamics in subjects with type 1 diabetes, the model output $y(t)$ [mmol/l] represents the deviation of glycemia $G(t)$ [mmol/l] from its basal value, i.e., the steady-state value $G_b$ [mmol/l]. Likewise, the first input $u(t)$ [U/min] denotes the deviation of the insulin administration rate from the basal insulin dosing rate $u_b$ [U/min]. The second input $d(t)$ [g/min] stands for the carbohydrate intake rate. The proposed nonparametric model defines the output $y(t)$ as

$$y(t) = \int_0^\infty g^u(\lambda)u(t-\lambda)\mathrm{d}\lambda + \int_0^\infty g^d(\lambda)d(t-\lambda)\mathrm{d}\lambda + \epsilon(t)\,, \quad (1)$$

where $g^u(t)$ is the impulse function of the insulin administration effect, $g^d(t)$ is the impulse function of the carbohydrate intake effect, both representing the convolution kernels.

The stochastic term $\epsilon(t) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ stands for the uncorrelated zero-mean random process, which reflects the continuous glucose monitoring sensor noise [27], [28], as well as the effects of various unmeasurable disturbances.

## III. BASIC ALGORITHM

The problem of estimating the impulse functions $g^u(t)$ and $g^d(t)$ of the model (1) is also called the deconvolution problem. To derive the correlation-based identification method, the cross-correlation function has to be introduced. The cross-correlation function $R_{xz}(\tau)$ of two general continuous-time infinite-length signals $x(t)$, $z(t)$ is, under the assumption of signal ergodicity, defined by the following integral transform for the lag argument $\tau \in \mathbb{R}$ [29]

$$R_{xz}(\tau) = E\{x(t+\tau)z(t)\} = \lim_{\vartheta\to\infty} \frac{1}{2\vartheta}\int_{-\vartheta}^{\vartheta} x(t+\tau)z(t)\mathrm{d}t\,. \quad (2)$$

Recall that if $x(t) = z(t)$, then $R_{xx}(\tau)$ is called the autocorrelation function.

Actually, the cross-correlation functions of the model output $y(t)$ with the inputs $u(t)$ and $d(t)$ are essential ones for our concern. The further presented equations result in a generalization of the Wiener-Hopf equation for two-input single-output systems, which was originally derived only for the single-input case, as can be found in the literature [7], [30]. According to equation (2), the cross-correlation function $R_{yu}(\tau)$ can be derived as

$$R_{yu}(\tau) = \lim_{\vartheta\to\infty} \frac{1}{2\vartheta}\int_{-\vartheta}^{\vartheta} y(t+\tau)u(t)\,\mathrm{d}t\,. \quad (3)$$

Substituting $y(t)$ from (1) into (3) yields

$$R_{yu}(\tau) = \lim_{\vartheta\to\infty} \frac{1}{2\vartheta}\int_{-\vartheta}^{\vartheta}\left[\int_0^\infty g^u(\lambda)u(t+\tau-\lambda)\mathrm{d}\lambda + \int_0^\infty g^d(\lambda)d(t+\tau-\lambda)\,\mathrm{d}\lambda + \epsilon(t+\tau)\right]u(t)\,\mathrm{d}t\,. \quad (4)$$

By changing the integral order, one can rewrite (4) to

$$R_{yu}(\tau) = \int_0^\infty g^u(\lambda)\lim_{\vartheta\to\infty}\left[\frac{1}{2\vartheta}\int_{-\vartheta}^{\vartheta} u(t+\tau-\lambda)u(t)\mathrm{d}t\right]\mathrm{d}\lambda + \int_0^\infty g^d(\lambda)\lim_{\vartheta\to\infty}\left[\frac{1}{2\vartheta}\int_{-\vartheta}^{\vartheta} d(t+\tau-\lambda)u(t)\mathrm{d}t\right]\mathrm{d}\lambda + \lim_{\vartheta\to\infty}\frac{1}{2\vartheta}\int_{-\vartheta}^{\vartheta}\epsilon(t+\tau)u(t)\mathrm{d}t\,. \quad (5)$$

Substituting the inner integrals of (5) as the autocorrelation function $R_{uu}(\tau-\lambda)$ and the cross-correlation function $R_{du}(\tau-\lambda)$, and the last integral as the cross-correlation function $R_{\epsilon u}(\tau)$, yields the generalized form of the Wiener–Hopf equation

$$R_{yu}(\tau) = \int_0^\infty g^u(\lambda)R_{uu}(\tau-\lambda)\,\mathrm{d}\lambda + \int_0^\infty g^d(\lambda)R_{du}(\tau-\lambda)\,\mathrm{d}\lambda + R_{\epsilon u}(\tau)\,. \quad (6)$$

Taking analogous steps, we derived the cross-correlation function $R_{yd}(\tau)$ for the second input as

$$R_{yd}(\tau) = \int_0^\infty g^d(\lambda)R_{dd}(\tau-\lambda)\,\mathrm{d}\lambda + \int_0^\infty g^u(\lambda)R_{ud}(\tau-\lambda)\,\mathrm{d}\lambda + R_{\epsilon d}(\tau)\,. \quad (7)$$

Notice the apparent symmetry in equations (6) and (7).

Since the signal $\epsilon(t)$ is supposed to be the uncorrelated noise, the cross-correlation functions $R_{\epsilon u}(\tau)$, $R_{\epsilon d}(\tau)$ in equations (6), (7) are simply zero, but only under the assumption of infinite-length experiment and the noise ergodicity. However, for a real experiment, a finite observation time $\vartheta \neq \infty$ is assumed instead. Consequently, all the cross-correlation and

the autocorrelation functions in (6) and (7) have to be replaced with their estimates $\tilde{R}_{uu}(\tau)$, $\tilde{R}_{dd}(\tau)$, $\tilde{R}_{ud}(\tau)$, $\tilde{R}_{du}(\tau)$, $\tilde{R}_{yu}(\tau)$, $\tilde{R}_{yd}(\tau)$ respectively [29]

$$\tilde{R}_{xz}(\tau) = \frac{1}{\vartheta - \tau} \int_0^{\vartheta-\tau} x(t+\tau)z(t)\,\mathrm{d}t\ . \tag{8}$$

Needless to say that the cross-correlation functions $\tilde{R}_{\epsilon u}(\tau)$, $\tilde{R}_{\epsilon d}(\tau)$ are no longer zero in this case, so equations (6), (7) will preserve their stochastic character.

## IV. DISCRETE-TIME FORM

For practical data-driven identification, the input and output signals are sampled uniformly with the sample time $T_s$ and therefore have a discrete-time nature. If $x(t)$ is a general continuous-time signal, then we introduce the notation $x_{(k)} = x(kT_s)$ where $k \in \mathbb{N}$ represents the sample index.

Moreover, we relate the coefficients of the discrete-time impulse function $g_i$ to the continuous-time impulse function $g(t)$ such that

$$g_i = g(iT_s)T_s\ . \tag{9}$$

Accordingly, the continuous-time model (1) will be transformed into the corresponding discrete-time form. The convolution integrals in (1) can be approximated by the finite summations, while the infinitesimal element $\mathrm{d}\lambda$ is replaced by $T_s > 0$, which is being absorbed into $g_n$ according to (9), yielding the finite impulse response model

$$y_{(k)} = \sum_{i=0}^{M_u} g_i^u u_{(k-i)} + \sum_{i=0}^{M_d} g_i^d d_{(k-i)} + \epsilon_{(k)}\ , \tag{10}$$

where $M_u$ and $M_d$ are the assumed lengths of the impulse response coefficients vectors $g^u$ and $g^d$, respectively.

Similarly, the integral in the correlation function (8) can be approximated by the finite summation [29], [31]

$$\tilde{R}_{xz}(nT_s) \approx \hat{R}_{xz}(n) = \frac{T_s}{N-n} \sum_{k=1}^{N-n} x_{(k+n)}z_{(k)}\ , \tag{11}$$

where $N$ denotes the number of samples of the processed time series $x$, $z$ and $n \in \mathbb{Z}$ is the integer lag argument satisfying the condition $n < N$. In addition, the symmetry property

$$\hat{R}_{xz}(-n) = \hat{R}_{zx}(n) \tag{12}$$

holds for (11) [29]. The discrete-time form of the generalized Wiener-Hopf equations (6) and (7) can be derived as

$$\hat{R}_{yu}(n) = \sum_{i=0}^{M_u} g_i^u \hat{R}_{uu}(n-i) + \sum_{i=0}^{M_d} g_i^d \hat{R}_{du}(n-i) + \hat{R}_{\epsilon u}(n)\ , \tag{13}$$

$$\hat{R}_{yd}(n) = \sum_{i=0}^{M_d} g_i^d \hat{R}_{dd}(n-i) + \sum_{i=0}^{M_u} g_i^u \hat{R}_{ud}(n-i) + \hat{R}_{\epsilon d}(n)\ , \tag{14}$$

where $n = 0 \ldots P$ is the lag argument. Note, that the maximum lag number $P$ should satisfy the condition

$$P \ll N\ . \tag{15}$$

### A. STATISTICAL PROPERTIES OF CROSS-CORRELATION FUNCTIONS

The cross-correlation functions $\hat{R}_{\epsilon u}(n)$, $\hat{R}_{\epsilon d}(n)$ in equations (13), (14) can not be directly estimated in practice because the noise term $\epsilon$ is unmeasurable. However, we may at least analyze the statistical properties of these cross-correlation functions. To this end, we will introduce the random vectors $\zeta^{\epsilon u} \in \mathbb{R}^{P+1\times 1}$, $\zeta^{\epsilon d} \in \mathbb{R}^{P+1\times 1}$ comprising the theoretical $\hat{R}_{\epsilon u}(n)$, $\hat{R}_{\epsilon d}(n)$ for different values of the argument $n$ as

$$\zeta^{\epsilon u} = \left[\hat{R}_{\epsilon u}(0)\ \hat{R}_{\epsilon u}(1)\ \ldots\ \hat{R}_{\epsilon u}(P)\right]^{\mathrm{T}}\ , \tag{16a}$$

$$\zeta^{\epsilon d} = \left[\hat{R}_{\epsilon d}(0)\ \hat{R}_{\epsilon d}(1)\ \ldots\ \hat{R}_{\epsilon d}(P)\right]^{\mathrm{T}}\ . \tag{16b}$$

The vectors $\zeta^{\epsilon u}$ and $\zeta^{\epsilon d}$ can be joint into

$$\zeta = \begin{bmatrix} \zeta^{\epsilon u} \\ \zeta^{\epsilon d} \end{bmatrix}\ . \tag{17}$$

Taking the expectancy operator to $\hat{R}_{\epsilon u}(n)$ in the terms of equation (11) yields

$$\begin{aligned} E\left\{\hat{R}_{\epsilon u}(n)\right\} &= \frac{T_s}{N-n} E\left\{\sum_{k=1}^{N-n} \epsilon_{(k+n)}u_{(k)}\right\} \\ &= \frac{T_s}{N-n} \sum_{k=1}^{N-n} E\left\{\epsilon_{(k+n)}\right\} u_{(k)} = 0\ . \end{aligned} \tag{18}$$

Note, that the above property holds also for $\hat{R}_{\epsilon d}(n)$. With regard to this finding, it can be deduced that the mean of the vector $\zeta$ defined by (17) is the zero vector, hence $E\{\zeta\} = \mathbf{0}$.

The covariance matrix $\mathcal{Q} \in \mathbb{R}^{2(P+1)\times 2(P+1)}$ of the vector $\zeta$ can be divided into four block submatrices and is defined as

$$\mathcal{Q} = \begin{pmatrix} \mathcal{Q}^{\epsilon\epsilon uu} & \mathcal{Q}^{\epsilon\epsilon ud} \\ \mathcal{Q}^{\epsilon\epsilon du} & \mathcal{Q}^{\epsilon\epsilon dd} \end{pmatrix} = E\left\{\begin{pmatrix} \zeta^{\epsilon u}\zeta^{\epsilon u\mathrm{T}} & \zeta^{\epsilon u}\zeta^{\epsilon d\mathrm{T}} \\ \zeta^{\epsilon d}\zeta^{\epsilon u\mathrm{T}} & \zeta^{\epsilon d}\zeta^{\epsilon d\mathrm{T}} \end{pmatrix}\right\} \tag{19}$$

The $i$-th row and the $j$-th column element of the covariance matrix $\mathcal{Q}^{\epsilon\epsilon ud} \in \mathbb{R}^{P+1\times P+1}$ can be derived according to the correlation function estimate (11) and the definitions (16a), (16b) of the vectors $\zeta^{\epsilon u}$, $\zeta^{\epsilon d}$ as

$$\begin{aligned} \mathcal{Q}_{ij}^{\epsilon\epsilon ud} &= E\left\{\hat{R}_{\epsilon u}(i)\hat{R}_{\epsilon d}(j)\right\} \\ &= \frac{{T_s}^2}{(N-i)\,(N-j)} E\left\{\sum_{k=1}^{N-i} \epsilon_{(k+i)}u_{(k)} \sum_{l=1}^{N-j} \epsilon_{(l+j)}d_{(l)}\right\}\ . \end{aligned} \tag{20}$$

By customizing a formula that can be found in [32], the following identity for two general random vectors $X$ and $Y$ is crucial at this point.

$$E\left\{\sum_{k=1}^{m} X_k \sum_{l=1}^{n} Y_l\right\} = \sum_{k=1}^{m}\sum_{l=1}^{n} E\{X_k Y_l\} \tag{21}$$

Applying the above identity, equation (20) reduces to

$$\mathcal{Q}_{ij}^{\epsilon\epsilon du} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sum_{k=1}^{N-i}\sum_{l=1}^{N-j} E\left\{\epsilon_{(k+i)}u_{(k)}\epsilon_{(l+j)}d_{(l)}\right\}\ . \tag{22}$$

Because $u$ and $d$ are deterministic, equation (22) can be rearranged as

$$\mathcal{Q}_{ij}^{\epsilon\epsilon ud} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sum_{k=1}^{N-i} \sum_{l=1}^{N-j} E\left\{\epsilon_{(k+i)}\epsilon_{(l+j)}\right\} u_{(k)} d_{(l)} \,. \tag{23}$$

According to the definition (2), assuming $E\left\{\epsilon_{(k+i)}\epsilon_{(l+j)}\right\} = R_{\epsilon\epsilon}(k+i-l-j)$ yields

$$\mathcal{Q}_{ij}^{\epsilon\epsilon ud} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sum_{k=1}^{N-i} \sum_{l=1}^{N-j} R_{\epsilon\epsilon}(k+i-l-j) u_{(k)} d_{(l)} \,. \tag{24}$$

Since $\epsilon$ is considered uncorrelated noise, its autocorrelation function is the Dirac delta function

$$R_{\epsilon\epsilon}(w) = \begin{cases} \sigma_\epsilon^2 & w = 0 \\ 0 & w \neq 0 \end{cases} \tag{25}$$

with the argument $w = k+i-l-j$ and the measure equal to the noise variance $\sigma_\epsilon^2$ [29], [33]. Thanks to this property, the double summation in (24) $\forall i \geq j$ simplifies to

$$\mathcal{Q}_{ij}^{\epsilon\epsilon ud} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sigma_\epsilon^2 \sum_{k=1}^{N-i} u_{(k)} d_{(k+i-j)} \,. \tag{26}$$

In the case $i < j$ complementary to (26), the formula can be obtained as

$$\mathcal{Q}_{ij}^{\epsilon\epsilon ud} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sigma_\epsilon^2 \sum_{l=1}^{N-j} d_{(l)} u_{(l+j-i)} \,. \tag{27}$$

One may notice that $\mathcal{Q}_{ji}^{\epsilon\epsilon ud} = \mathcal{Q}_{ij}^{\epsilon\epsilon du}$ what implies that $\mathcal{Q}^{\epsilon\epsilon du} = \left(\mathcal{Q}^{\epsilon\epsilon ud}\right)^{\mathrm{T}}$. The above steps can be taken to derive the remaining submatrices $\mathcal{Q}^{\epsilon\epsilon uu}$, $\mathcal{Q}^{\epsilon\epsilon dd}$ of (19) yielding the formulas

$$\mathcal{Q}_{ij}^{\epsilon\epsilon uu} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sigma_\epsilon^2 \sum_{k=1}^{N-i} u_{(k)} u_{(k+i-j)} \,, \tag{28}$$

$$\mathcal{Q}_{ij}^{\epsilon\epsilon dd} = \frac{{T_s}^2}{(N-i)\,(N-j)} \sigma_\epsilon^2 \sum_{k=1}^{N-i} d_{(k)} d_{(k+i-j)} \,. \tag{29}$$

Since the covariance matrix (19) is symmetric, the submatrices $\mathcal{Q}^{\epsilon\epsilon uu}$, $\mathcal{Q}^{\epsilon\epsilon dd}$ are also symmetric, implying that $\mathcal{Q}_{ji}^{\epsilon\epsilon uu} = \mathcal{Q}_{ij}^{\epsilon\epsilon uu}$ and $\mathcal{Q}_{ji}^{\epsilon\epsilon dd} = \mathcal{Q}_{ij}^{\epsilon\epsilon dd}$.

Notice that increasing the lag argument $n$ in estimate (11) truncates the summation interval to $N-n$, what yields its higher variance, so this estimate should be considered less accurate and confident.

## V. ESTIMATE OF IMPULSE RESPONSE COEFFICIENTS

For estimating the impulse response coefficients of the nonparametric model (10) we will utilize the generalized least squares method [33], [34] applied to the derived discrete-time Wiener-Hopf equations (13), (14). The parameter vector $g \in \mathbb{R}^{M_u+M_d+2\times 1}$ can be formally noted as

$$g = \begin{bmatrix} g^u \\ g^d \end{bmatrix} \,, \tag{30}$$

where the subvectors $g^u \in \mathbb{R}^{M_u+1\times 1}$ and $g^d \in \mathbb{R}^{M_d+1\times 1}$ are defined as

$$g^u = \begin{bmatrix} g_0^u & g_1^u & g_2^u & \cdots & g_{M_u}^u \end{bmatrix}^{\mathrm{T}} \,, \tag{31a}$$

$$g^d = \begin{bmatrix} g_0^d & g_1^d & g_2^d & \cdots & g_{M_d}^d \end{bmatrix}^{\mathrm{T}} \,. \tag{31b}$$

The cross-correlation functions $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$ from the Wiener-Hopf equations (13),(14), respectively, have to be reshaped into the equivalent matrix form (32), as shown at the bottom of the page, assuming the lag argument $n = 0 \ldots P$. The symmetry property (12) was also accounted in (32). The equation system (32) is overdetermined if condition

$$M_u + M_d < 2P \tag{33}$$

holds, while condition (15) must be satisfied as well.

The choice of the maximum lag number $P$ should be based primarily on the character of the obtained cross-correlation functions $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$. If these cross-correlation functions converge at some $n$, then increasing $P$ over this value typically does not have any positive impact on the estimate quality. Also, the number of processed samples $N$ should be taken into account, since increasing the maximal lag argument $P$ truncates the summation interval in the estimate (11) to $N-P$, thus making it less accurate and highly uncertain. It is a good practice for $P$ to not exceed $\frac{N}{2}$.

Concerning the lengths $M_u$, $M_d$ of the estimated response coefficients vectors, their choice should reflect the sample

$$\begin{pmatrix} \hat{R}_{yu}(0) \\ \hat{R}_{yu}(1) \\ \vdots \\ \hat{R}_{yu}(P) \\ \hat{R}_{yd}(0) \\ \hat{R}_{yd}(1) \\ \vdots \\ \hat{R}_{yd}(P) \end{pmatrix} = \begin{pmatrix} \hat{R}_{uu}(0) & \hat{R}_{uu}(1) & \ldots & \hat{R}_{uu}(M_u) & \hat{R}_{du}(0) & \hat{R}_{ud}(1) & \ldots & \hat{R}_{ud}(M_d) \\ \hat{R}_{uu}(1) & \hat{R}_{uu}(0) & \ldots & \hat{R}_{uu}(M_u-1) & \hat{R}_{du}(1) & \hat{R}_{du}(0) & \ldots & \hat{R}_{ud}(M_d-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{uu}(P) & \hat{R}_{uu}(P-1) & \ldots & \hat{R}_{uu}(P-M_u) & \hat{R}_{du}(P) & \hat{R}_{du}(P-1) & \ldots & \hat{R}_{du}(P-M_d) \\ \hat{R}_{ud}(0) & \hat{R}_{du}(1) & \ldots & \hat{R}_{du}(M_u) & \hat{R}_{dd}(0) & \hat{R}_{dd}(1) & \ldots & \hat{R}_{dd}(M_d) \\ \hat{R}_{ud}(1) & \hat{R}_{ud}(0) & \ldots & \hat{R}_{du}(M_u-1) & \hat{R}_{dd}(1) & \hat{R}_{dd}(0) & \ldots & \hat{R}_{dd}(M_d-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{ud}(P) & \hat{R}_{ud}(P-1) & \ldots & \hat{R}_{ud}(P-M_u) & \hat{R}_{dd}(P) & \hat{R}_{dd}(P-1) & \ldots & \hat{R}_{dd}(P-M_d) \end{pmatrix} \begin{pmatrix} g_0^u \\ g_1^u \\ \vdots \\ g_{M_u}^u \\ g_0^d \\ g_1^d \\ \vdots \\ g_{M_d}^d \end{pmatrix} + \begin{pmatrix} \hat{R}_{\epsilon u}(0) \\ \hat{R}_{\epsilon u}(1) \\ \vdots \\ \hat{R}_{\epsilon u}(P) \\ \hat{R}_{\epsilon d}(0) \\ \hat{R}_{\epsilon d}(1) \\ \vdots \\ \hat{R}_{\epsilon d}(P) \end{pmatrix} \tag{32}$$

time and the anticipated dynamics of the system identified while covering the dominant part of the impulse response. We recommend adjusting $M$ such that the last estimated element $\hat{g}_M$ is below 10% of the peak value of the estimated impulse response.

The full equation system (32) can be written in a compact form

$$\begin{pmatrix} \hat{\mathcal{R}}_{yu} \\ \hat{\mathcal{R}}_{yd} \end{pmatrix} = \begin{pmatrix} \hat{\mathcal{R}}_{uu} & \hat{\mathcal{R}}_{du} \\ \hat{\mathcal{R}}_{ud} & \hat{\mathcal{R}}_{dd} \end{pmatrix} \begin{pmatrix} g^u \\ g^d \end{pmatrix} + \begin{pmatrix} \zeta^{\epsilon u} \\ \zeta^{\epsilon d} \end{pmatrix} , \tag{34}$$

and even more simplified as

$$\hat{\mathcal{R}}_{\mathcal{Y}} = \hat{\mathcal{R}}_{\mathcal{U}} g + \zeta \, . \tag{35}$$

The structures of submatrices $\hat{\mathcal{R}}_{uu} \in \mathbb{R}^{P+1\times M_u+1}$, $\hat{\mathcal{R}}_{dd} \in \mathbb{R}^{P+1\times M_d+1}$, $\hat{\mathcal{R}}_{ud} \in \mathbb{R}^{P+1\times M_u+1}$, $\hat{\mathcal{R}}_{du} \in \mathbb{R}^{P+1\times M_d+1}$ and vectors $\hat{\mathcal{R}}_{yu} \in \mathbb{R}^{P+1\times 1}$, $\hat{\mathcal{R}}_{yd} \in \mathbb{R}^{P+1\times 1}$ in (34), as well as matrix $\hat{\mathcal{R}}_{\mathcal{U}} \in \mathbb{R}^{2(P+1)\times M_u+M_d+2}$ and vector $\hat{\mathcal{R}}_{\mathcal{Y}} \in \mathbb{R}^{2(P+1)\times 1}$ in (35) result from the full equation system (32), while the random vectors $\zeta^{\epsilon u}$, $\zeta^{\epsilon d}$ and $\zeta$ were defined in (17).

It is necessary to say that after some detailed investigation, we found some structural similarities between the equation system (32) and the equations published in [35], yet the authors have taken different steps within its derivation and also fully neglected the stochastic nature of the cross-correlation functions. Moreover, the equation system presented in the aforementioned paper was not proposed to be solved in the least squares sense but as a regular linear equation system, so neither the regularization strategies were applied.

The following assumption needs to be added if one wants to study the method presented in [35]

$$M_u + M_d = 2P \, . \tag{36}$$

The above assumption implies that the length of the vector $\hat{\mathcal{R}}_{\mathcal{Y}} \in \mathbb{R}^{M_u+M_d+2\times 1}$ is equal to the number of estimated parameters and $\hat{\mathcal{R}}_{\mathcal{U}} \in \mathbb{R}^{M_u+M_d+2\times M_u+M_d+2}$ is a square matrix, so the parameter estimate can be obtained as the solution of the regular linear equation system (35)

$$\hat{g} = \hat{\mathcal{R}}_{\mathcal{U}}^{-1} \hat{\mathcal{R}}_{\mathcal{Y}} \, . \tag{37}$$

On the other hand, in [36] a similar correlation method was generalized for the multiple-input multiple-output system but under the theoretical assumption of random uncorrelated input signals. Although this assumption significantly simplified problem formulation and solution, it automatically hampered the application of this method to identify empirical models of diabetes, as using random sequences to excite biological systems is simply unfeasible or extremely dangerous.

For the least squares-based estimation of the impulse response coefficients vector $\hat{g}$, the residuals vector $e \in \mathbb{R}^{2(P+1)\times 1}$ has to be introduced as

$$e = \begin{bmatrix} e^{yu} \\ e^{yd} \end{bmatrix} = \hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}} \hat{g} = \begin{bmatrix} \hat{\mathcal{R}}_{yu} - \hat{\mathcal{R}}_{uu}\hat{g}^u - \hat{\mathcal{R}}_{du}\hat{g}^d \\ \hat{\mathcal{R}}_{yd} - \hat{\mathcal{R}}_{dd}\hat{g}^d - \hat{\mathcal{R}}_{ud}\hat{g}^u \end{bmatrix} . \tag{38}$$

The cost function of the generalized least squares method modified by adding the regularization of the estimate is defined by the quadratic form

$$J(\hat{g}) = \frac{1}{2} \left[ \left( \hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}} \hat{g} \right)^{\mathrm{T}} \mathcal{Q}^{-1} \left( \hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}} \hat{g} \right) + \hat{g}^{\mathrm{T}} \Lambda \hat{g} \right] \tag{39}$$

with respect to the estimated parameter vector $\hat{g}$. In the above equation, $\mathcal{Q}$ is the covariance matrix of the noise vector $\zeta$ and $\Lambda \in \mathbb{R}^{M_u+M_d+2\times M_u+M_d+2}$ is a positive-definite symmetric regularization matrix, the design of which shall be clarified later.

An intuitive rationale for using the generalized least squares method is that it induces a "weighting" and "decorrelating" effect on the residuals by the square root of the inverse covariance matrix $\mathcal{Q}^{-1}$ of the noise vector $\zeta$. If the variance of the noise in the linear regression problem is not uniform, then samples with larger variance should be less trusted when calculating the parameter estimate than those with lower variance.

The gradient of the cost function (39) with respect to the estimated vector $\hat{g}$ can be derived as

$$\nabla_{\hat{g}} J(\hat{g}) = -\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \left( \hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}} \hat{g} \right) + \Lambda \hat{g} \, . \tag{40}$$

According to the optimality condition $\nabla_{\hat{g}} J(\hat{g}) = \mathbf{0}$, the optimal parameter estimate $\hat{g}$ can be obtained in a closed form

$$\hat{g} = \left( \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda \right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{Y}} \, . \tag{41}$$

One of the emerging advantages of the presented method over traditional prediction error methods is that the computational complexity and dimensions of the problem do not depend on the number of processed samples $N$.

### A. ESTIMATE REGULARIZATION

Regularization is applied in order to involve some prior knowledge of the system being identified in the estimate and also to lower the estimate variance. Ultimately, the regularization term $\hat{g}^{\mathrm{T}} \Lambda \hat{g}$ in the cost function (39) acts as a penalty for specific characteristics of the system [37] or can be seen as penalizing certain parameters describing "unlikely" systems [38]. It should be noted that the reduction in the variance of the parameter estimate induced by regularization comes at the expense of an emerging bias, so a bias–variance trade-off must always be made [39].

One way of involving the prior knowledge within the regularization is via the covariance matrix of the system parameters [39], yet this strategy obviously requires to ever have this prior knowledge, which may be unfeasible in many applications, so in this paper we will rather choose heuristic regularization strategies.

The particular effect of regularization depends on the linear transform operator $L \in \mathbb{R}^{n_\Gamma \times M_u+M_d+2}$ and the diagonal scaling matrix $\Gamma \in \mathbb{R}^{n_\Gamma \times n_\Gamma}$, so one can theoretically decompose the matrix $\Lambda$ as [37]

$$\hat{g}^{\mathrm{T}} \Lambda \hat{g} = \left( L\hat{g} \right)^{\mathrm{T}} \Gamma L \hat{g} = \hat{g}^{\mathrm{T}} L^{\mathrm{T}} \Gamma L \hat{g} \, , \tag{42}$$

where $\Gamma$ is a diagonal scaling matrix

$$\Gamma = \mathrm{diag}(\gamma)\,. \tag{43}$$

If multiple types of penalty are combined, the regularization matrix $\Lambda$ results from the sum $\Lambda = \Lambda^A + \Lambda^B + \Lambda^C$. In the framework of this paper, three penalties are assumed. In particular, the smoothing operation and regularization to provide asymptotic stability and causality are applied to the estimated impulse response coefficients vector $\hat{g}$.

The aforementioned smoothing effect can be achieved by penalizing the sum of the squared differences of the estimated impulse response coefficients vector $\hat{g}$. In the frequency domain, this operation can be interpreted as low-pass filtering of the impulse function $\hat{g}_n$ [37], since the penalty kernel $\mathbf{h} = \begin{pmatrix}1 & -1\end{pmatrix}$ represents a simple high-pass filter. The corresponding linear transform $L^A \in \mathbb{R}^{M_u+M_d \times M_u+M_d+2}$ can be formed as [37]

$$L^A = \begin{pmatrix} L^s & \mathbf{0} \\ \mathbf{0} & L^s \end{pmatrix} \quad L^s = \mathrm{diag}\left(\mathbf{h}, \ldots, \mathbf{h}\right)\,. \tag{44}$$

However, the undesired side effect of this regularization is the suppression of the typical peak of the impulse function. Since there is a general expectation of a stable impulse response with typically larger changes at the beginning and a relatively slow exponential decay character, the scaling vector of the regularization matrix $\Lambda^A$ may comprise the coefficients of a geometric series

$$\gamma_i^{(\star)} = a\, q^{i-1} \text{ for } u \to \star \text{ and } d \to \star \tag{45}$$

with $q > 1$ and $10^{-3} < a < 10^2$. In this way, the regularization-induced distortion of the impulse function could be minimized while retaining the desired smoothing effect. The parameter $a$ proportionally affects the overall strength of the smoothing effect, but choosing too high values typically causes a significant deformation of the impulse response and hence induces an unacceptable bias, so this parameter should be adjusted carefully. The coefficient $q$ should be chosen with regard to the anticipated dynamics of the identified system, while the larger $q$ corresponds to faster dynamics, and the smaller $q$ is better for slower systems.

The second regularization strategy is proposed to ensure the asymptotic stability of the identified model. In general, the continuous-time impulse response model (1) is stable if $\lim_{t\to\infty} g(t) = 0$. However, the discrete-time finite impulse response model (10) is structurally stable, even if $g_M \neq 0$, but its asymptotic stability $\lim_{M\to\infty} g_M = 0$ must be guaranteed anyway. Consequently, a quadratic penalty can be used for this purpose, while the linear transform $L^B$ is equal to the unit matrix $L^B = I$, and the scaling vector $\gamma^B$ may comprise elements of a geometric series (45) with $q > 1$ and $10^{-3} < a < 10^2$ in order to reflect the exponential decay nature of the impulse response of a stable aperiodic system [39]. The tuning rules for the coefficients $a$ and $q$ are similar to those proposed for the smoothing regularization.

The last regularization strategy issues the causality of the identified model (10), since for a strictly causal model $g_0 = 0$ is demanded. Instead of implementing it as a hard constraint, a dedicated quadratic penalty may be applied. The transform matrix will be simply $L^C = I$, and the scaling vector $\gamma^C$ should comprise a large value at the positions corresponding to $\hat{g}_0$.

### B. PARAMETER ESTIMATE UNCERTAINTIES

In addition to the optimal parameter estimate $\hat{g}$ determined according to (41), the confidence interval that represents the parameter uncertainties may also be useful to assess the identification results.

The covariance matrix $\mathcal{P} \in \mathbb{R}^{M_u+M_d+2 \times M_u+M_d+2}$ of the estimate $\hat{g}$ can be derived starting from the definition

$$\mathcal{P} = \begin{pmatrix} \mathcal{P}^u & \mathcal{P}^{ud} \\ \mathcal{P}^{du} & \mathcal{P}^d \end{pmatrix} = E\left\{\left(\hat{g} - E\left\{\hat{g}\right\}\right)\left(\hat{g} - E\left\{\hat{g}\right\}\right)^{\mathrm{T}}\right\}\,. \tag{46}$$

Applying the expectancy operator to the estimate (41) while substituting $\hat{\mathcal{R}}_{\mathcal{Y}}$ from (35) results into

$$E\left\{\hat{g}\right\} = \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \left(\hat{\mathcal{R}}_{\mathcal{U}} g + E\left\{\zeta\right\}\right). \tag{47}$$

As was shown in (18), the noise vector $\zeta$ is zero-mean, what yields

$$E\left\{\hat{g}\right\} = \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} g\,. \tag{48}$$

The above equation implies that the regularization makes the estimate biased, so conversely, in the limit case of (48) $\lim_{\Lambda\to\mathbf{0}} E\left\{\hat{g}\right\} = g$ the estimate would be unbiased.

Continuing deriving the covariance matrix $\mathcal{P}$ as outlined in (46), the difference $\hat{g} - E\left\{\hat{g}\right\}$ can be substituted according to (35), (41) and (48) as

$$\hat{g} - E\left\{\hat{g}\right\} = \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \zeta\,, \tag{49}$$

what along with the covariance matrix $\mathcal{Q}$ of the noise vector $\zeta$ from (19) yields (50), as shown at the bottom of the next page. Thanks to the generalized least squares method, the covariance matrix $\mathcal{P}$ is minimal [34], [33] and its formula is

$$\begin{aligned}\mathcal{P} = &\left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \\ &\times \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1}\end{aligned} \tag{51}$$

Concerning the confidence interval of $\hat{g}$, one may assume the standard normally distributed random vector $r$

$$r = \left(\hat{g} - E\left\{\hat{g}\right\}\right) \oslash \sqrt{\mathrm{diag}(\mathcal{P})}\,, \tag{52}$$

where $\sqrt{\mathrm{diag}(\mathcal{P})}$ is the element-wise square root of the diagonal vector of the matrix $\mathcal{P}$ and $\oslash$ denotes the element-wise division.

The $1-\alpha$ probability of (52) gets

$$1-\alpha = P\left(-r_{1-\frac{\alpha}{2}} \leq \left(\hat{g} - E\left\{\hat{g}\right\}\right) \oslash \sqrt{\mathrm{diag}(\mathcal{P})} \leq r_{1-\frac{\alpha}{2}}\right) \tag{53}$$

Substituting $E\left\{\hat{g}\right\}$ form (48) into (53) yields (54), as shown at the bottom of the page. Multiplying the inequality in (54) by $\Psi = I + \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}}\right)^{-1} \Lambda$ one gets the confidence interval $\left(\underline{\hat{g}}, \overline{\hat{g}}\right)$ as

$$\overline{\hat{g}} / \underline{\hat{g}} = \Psi \left( \hat{g} \pm \sqrt{\mathrm{diag}(\mathcal{P})} r_{1-\frac{\alpha}{2}} \right) , \tag{55}$$

whence the quantile $r_{1-\frac{\alpha}{2}}$ satisfies $P\left(r_{1-\frac{\alpha}{2}}\right) = 1 - \frac{\alpha}{2}$ of the cumulative distribution function $P$ of the standard normal distribution.

## VI. PARAMETRIC MODEL APPROXIMATION

The nonparametric discrete-time model (10) with the coefficients of impulse responses identified as $\hat{g}^u, \hat{g}^d$ can be approximated by the parametric transfer function-based model

$$y_{(k)} = \frac{B^u(z)}{A^u(z)} u_{(k)} + \frac{B^d(z)}{A^d(z)} d_{(k)} + \epsilon_{(k)} , \tag{56}$$

where the input and output signals are equivalent to those of the nonparametric model (1).

We may unwind the polynomials of the model (56) as

$$A^\star(z) = 1 + a_1^\star z^{-1} + \ldots a_{n_{A^\star}}^\star z^{-n_{A^\star}} \tag{57a}$$

$$B^\star(z) = b_1^\star z^{-1} + b_2^\star z^{-2} + \ldots b_{n_{B^\star}}^\star z^{-n_{B^\star}} , \tag{57b}$$

where $n_{A^u}, n_{B^u}$ denote the orders of the insulin administration effect submodel and $n_{A^d}, n_{B^d}$ are the orders of the carbohydrate intake effect submodel.

The strategy to be presented can be seen as the second phase of the full-featured two-step identification procedure for the transfer function-based model (56). The traditional direct approaches to identify the parametric model (56) include the instrumental variables method [7] or the numerical minimization of the prediction error [22].

The rationale for choosing the empirical model of glycemia dynamics (56) having different feedback (poles) for both inputs is based on the physiology of type 1 diabetes, since the administered insulin and the consumed carbohydrates are metabolized mostly in different body compartments [28], [40], so this assumption automatically excludes the ARX model.

To separately estimate the parameter vector $\theta^\star \in \mathbb{R}^{n_{B^\star}+n_{A^\star}\times 1}$ for each partial transfer function of the parametric model (56), the generalized least squares method can conveniently be adopted also in this case.

$$\theta^\star = \left[ b_1^\star \ \ldots \ b_{n_{B^\star}}^\star \ a_1^\star \ \ldots \ a_{n_{A^\star}}^\star \right]^{\mathrm{T}} \tag{58}$$

One can derive the linear regression system

$$H^\star \hat{\theta}^\star = \hat{g}^\star + \chi^\star \text{ for } u \to \star \text{ and } d \to \star , \tag{59}$$

where $\chi^\star \in \mathbb{R}^{M_\star+1\times 1}$ is the uncertainty of the impulse response estimate and the regression matrix $H^\star \in \mathbb{R}^{M_\star+1\times n_{B^\star}+n_{A^\star}}$ is defined as

$$H^\star = \begin{pmatrix} 1 & 0 & \ldots & 0 & -\hat{g}_0^\star & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & -\hat{g}_1^\star & -\hat{g}_0^\star & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 & -\hat{g}_{M_\star-1}^\star & -\hat{g}_{M_\star-2}^\star & \cdots & -\hat{g}_{M_\star-n_{A^\star}}^\star \end{pmatrix} \tag{60}$$

The cost function of the generalized least squares method for this problem gets [33], [34]

$$J\left(\hat{\theta}^\star\right) = \frac{1}{2} \left(\hat{g}^\star - H^\star \hat{\theta}^\star\right)^{\mathrm{T}} \mathcal{P}^{\star -1} \left(\hat{g}^\star - H^\star \hat{\theta}^\star\right) . \tag{61}$$

The optimal parameter estimate can be obtained as

$$\hat{\theta}^\star = \left(H^{\star\mathrm{T}} \mathcal{P}^{\star -1} H^\star\right)^{-1} H^{\star\mathrm{T}} \mathcal{P}^{\star -1} \hat{g}^\star , \tag{62}$$

where the covariance submatrices $\mathcal{P}^u$, $\mathcal{P}^d$ were defined in (46) and (51). The estimate (62) has to be evaluated separately for both vectors $\hat{g}^u$, $\hat{g}^d$ estimated from the first step and both submodels of (56).

The flow diagram summarizing the important steps and equations of the proposed complex identification algorithm is depicted in Fig. 1.

By a holistic comparison of the proposed method with other approaches, we summarized the following remarks. Concerning the prediction error methods for discrete-time models, the commonly used single step-ahead prediction error criterion basically cannot provide a valid model with good prediction performance, whereas the improved multiple step-ahead prediction error criterion requires iterative numeric optimization algorithms to be involved. Moreover, the compliance of the thus identified model with the basic

$$\begin{aligned} \mathcal{P} &= E\left\{ \left[ \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \zeta \right] \left[ \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \zeta \right]^{\mathrm{T}} \right\} \\ &= \left[ \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \right] E\left\{\zeta \zeta^{\mathrm{T}}\right\} \left[ \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \right]^{\mathrm{T}} \\ &= \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \mathcal{Q} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \end{aligned} \tag{50}$$

$$1 - \alpha = P\left( \hat{g} - \sqrt{\mathrm{diag}(\mathcal{P})} r_{1-\frac{\alpha}{2}} \leq \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} + \Lambda\right)^{-1} \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}} \mathcal{Q}^{-1} \hat{\mathcal{R}}_{\mathcal{U}} g \leq \hat{g} + \sqrt{\mathrm{diag}(\mathcal{P})} r_{1-\frac{\alpha}{2}} \right) \tag{54}$$

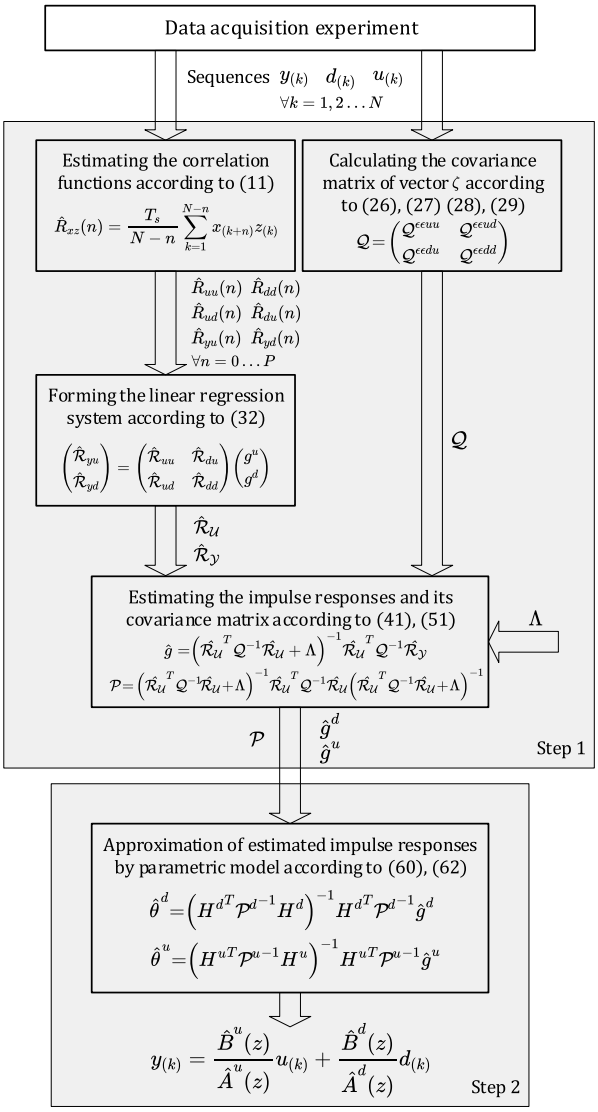



**FIGURE 1.** Flow diagram of the proposed identification algorithm.

physiology usually cannot be ensured under all circumstances. On the other hand, identification strategies for continuous-time models that are mostly based on minimizing the long-term model prediction error criterion lead to the formulation of nonlinear optimization problems that require special iterative solvers. Due to the correlation-based approach, the model identified using the proposed method should theoretically perform accurate long-term prediction of glycemia, while compliance with physiology can be ensured by applying a wide variety of regularization techniques. It is essential to mention the existence of an analytical closed-form solution of the parameter estimation problem that arose from the generalized least squares method formulation of the identification problem. Speaking of computational complexity, one of the nonparametric identification methods discussed in the literature review was based on a solution of the Tikhonov-type variational problem, so there was a lack of analytical solution also in this case. Comparing the proposed nonparametric method for estimating impulse responses to the full individualization of a complex physiology-based simulation model, it can be concluded that the proposed method requires only basic input-output data in the form of diabetic diary and continuous glucose measurements, whereas to identify a relatively large set of parameters of the simulation model, specialized clinical experiments involving radioactive tracers or even invasive measurements have to be used.

## VII. CASE STUDY

As outlined in the title, the target application domain of the proposed identification algorithm is the empirical model of glycemia dynamics in patients with type 1 diabetes. Empirical models provide a very simplified description of actually complex physiological phenomena that involve glucose kinetics, insulin-glucose interaction, insulin absorption, or carbohydrate ingestion. However, the main advantage of empirical models is that they can be individualized using only passive experiment diabetic data available from free-living conditions, namely the continuous glucose monitoring readings and the basic diabetic diary. In this section, virtual diabetic data generated by a complex physiology-based simulation model will be used for the validation of the proposed identification algorithm.

### A. DIABETIC DATA

A typical diabetic dataset comprising information on administered insulin boluses and carbohydrate intake in the form of a diabetic diary poses a specific challenge concerning the identifiability of empirical models. During conventional insulin therapy, the insulin dose is determined according to a simple rule called the bolus calculator. The bolus calculator utilizes the information on the current finger-stick glycemia measurement $G$ [mmol/l] and the expected carbohydrate content of the meal $CHO$ [g] to advise the compensating insulin bolus size $B$ [U] according to the following formula [41], [42]

$$B = \frac{G - G_w}{IS} + \frac{CHO}{ICR}, \tag{63}$$

where $IS$ [mmol/l/U] is the insulin sensitivity parameter, $ICR$ [g/U] is the insulin-carbohydrate ratio parameter and $G_w$ [mmol/l] is the target glycemia, usually chosen as $G_w = 5.5$ mmol/l.

Since the premeal bolus strategy is generally recommended by physicians [43], insulin is typically administered $t_B$ minutes before the corresponding meal intake [42]. The adverse characteristics of the traditional bolus calculator (63) is that the first input signal $u(t)$ is defined as a linear combination of the second input $d(t)$ and the output $y(t)$. Unfortunately, this linear dependence and the fixed meal-insulin timing cause both inputs $u(t)$ and $d(t)$ to be highly correlated and result in poor excitation quality and ultimately in an ill-conditioned identification problem [9]. A solution to this problem is to add a normally distributed random component $\eta \sim \mathcal{N}\left(0, \sigma_B^2\right)$ to the insulin bolus. Consequently, the bolus calculator (63)

can be reshaped as

$$u(t)=\begin{cases}\frac{y(t)+G_b-G_w}{IS}+\frac{d(t+t_B)}{ICR}+\eta(t) & d(t+t_B)\neq 0\\ 0 & \text{else}\end{cases}, \quad (64)$$

defining the insulin administration rate signal $u(t)$ using the signals $d(t)$, $y(t)$ featured in section II.

The glycemia response for this experiment was obtained in-silico, using the complex physiology-based nonlinear simulation model, discussed in [44] and [45]. The available mean population parameters reported in [44] were adopted, while some adjustments had to be made in order to simulate the metabolic specifics of type 1 diabetes [46]. The basal state of the model was calculated according to the basal glycemia $G_b = 6$ mmol/l and the corresponding basal insulin administration rate $u_b = 0.01$ U/min. It is worth noting that this model was accepted by the *Food and Drug Administration* agency as a substitute to animal trials for the preclinical testing of control strategies in artificial pancreas studies [47], so the results can be considered quite realistic and credible.

The data acquisition experiment was designed to mimic the regular insulin treatment of a type 1 diabetic subject during the 6-day period with an overall number of 25 meals and a total carbohydrate amount of 433 g. The virtual continuous glucose monitoring readings were sampled with the sample time $T_s = 20$ min and the total length of the experiment was $6\times 60\times 24$ min resulting in the number of samples $N = 433$. The glycemia measurements were distorted by the additive white noise with the standard deviation $\sigma_\epsilon = 0.1$ mmol/l.

The insulin treatment was executed according to the modified bolus calculator (64) with the parameters adjusted as $IS = 15$ mmol/l/U, $ICR = 8.0$ g/U and $t_B = 20$ min. The not less important parameter $\sigma_B^2$, representing the variance of the random bolus component, was chosen as $\sigma_B^2 = 0.05$ U since this value turned out to be a reasonable trade-off between system excitation and patient safety.

The resulting input-output identification diabetic dataset is depicted in Fig. 2 where one can see the noisy nature of the glycemia measurements as well as the time advance of insulin administrations before the corresponding meal intake impulses. It can also be observed that the ratio of the input impulse signals $u(t)/d(t)$ is variable, as this is highly desired and vital for ensuring the identifiability of multiple-input systems.

### B. RESULTS

Concerning the tuning of the identification algorithm, the assumed number of lags $P$ in the equation system (32) was chosen as $P = 170$ to satisfy condition (15). Based on rough expectations about physiological responses of the human body, the lengths of the identified impulse responses were chosen as $M_u = 50$, $M_d = 50$, while satisfying condition (33). The scaling vector $\gamma^A$ figuring in the regularization matrix $\Lambda^A$ will be formed by two geometric series (45) as defined in (65). The scaling vector $\gamma^B$ for the regularization matrix $\Lambda^B$ will also comprise the geometric series (45) as documented in (66).

$$\gamma^A = 0.5\left[1\ 1.05\ \ldots\ 1.05^{M_u-2}\ 1\ 1.06\ \ldots\ 1.06^{M_d-2}\right]^{\mathrm{T}} \quad (65)$$

$$\gamma^B = 0.1\left[1\ 1.02\ \ldots\ 1.02^{M_u-2}\ 1\ 1.07\ \ldots\ 1.07^{M_d-2}\right]^{\mathrm{T}} \quad (66)$$

The vector $\gamma^C$ for the regularization matrix $\Lambda^C$ was chosen as

$$\gamma^C = \left[5\ 0\ \ldots\ 0\ 5\ 0\ \ldots\ 0\right]^{\mathrm{T}}. \quad (67)$$

The resulting combined regularization matrix $\Lambda$ was calculated as the sum of partial regularization matrices, and its entries can be seen in Fig. 3 where one can notice the diagonal entries increasing with the index number, the presence of small negative off-diagonal entries induced by the smoothing regularization, and also a dominant value of the first diagonal element due to the causality regularization.

The inverse $\mathcal{Q}^{-1}$ of the covariance matrix of the noise vector $\zeta$ was determined from the available input signals $u_{(k)}$, $d_{(k)}$ of the identification dataset and its entries are visualized as a colormap in Fig. 4. In Fig. 4 one can see the matrix $\mathcal{Q}^{-1}$ having a quasi-diagonal structure implying that the effect of residuals weighting in the generalized least squares method is more prominent than the decorrelating effect. Moreover, the block submatrices corresponding to both submodels have significantly different magnitudes of the elements, and it can also be as that the diagonal entries are decreasing with the index number, what can be interpreted as that the estimates of the correlation functions with the higher argument are less trusted when estimating the impulse functions.

The impulse response coefficients vectors of the model (10) were estimated in terms of equation (41). In addition to the optimal solution $\hat{g}$, the confidence interval $\underline{\hat{g}}$, $\overline{\hat{g}}$ for $\alpha = 0.05$ was determined according to equation (55) while the obtained results are summarized in a graphical form in Fig. 5. The presented results show that both estimated impulse responses represent a stable system with the aperiodic nature, as well as that the coefficients keep their sign uniform, what overally makes the estimate physiology compliant. There can also be observed a slower and longer-lasting effect of insulin administration with a peak at approximately 300 minutes compared to a faster and shorter-lasting effect of carbohydrate intake, which peaked at approximately 200 minutes.

We decided to compare the performance of the proposed strategy that is based on solving the overdetermined equation system (32) as the linear regression problem in the least squares sense (39) to the approach based on direct solution (37) of regular linear equation system under the assumption (36). The estimate of impulse responses obtained according to the latter strategy adopted from the referred paper [35] is depicted in Fig. 6. There can be observed apparent distortion and noise in the estimated impulse responses, concluding that this relatively straightforward method cannot ensure general validity of estimate making it insufficient for our concern.

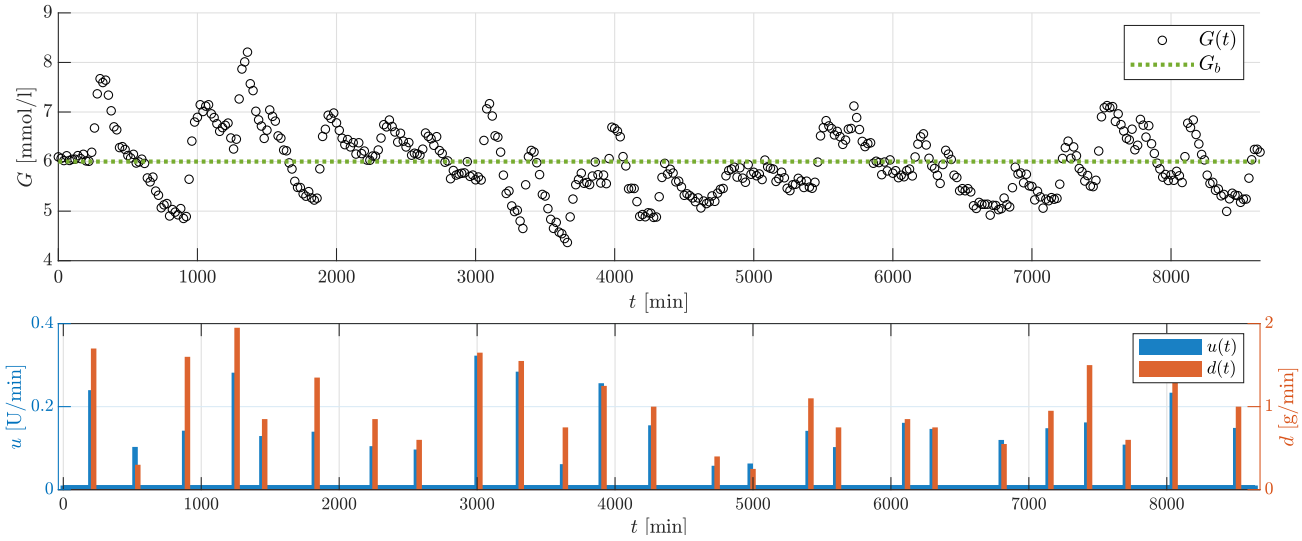
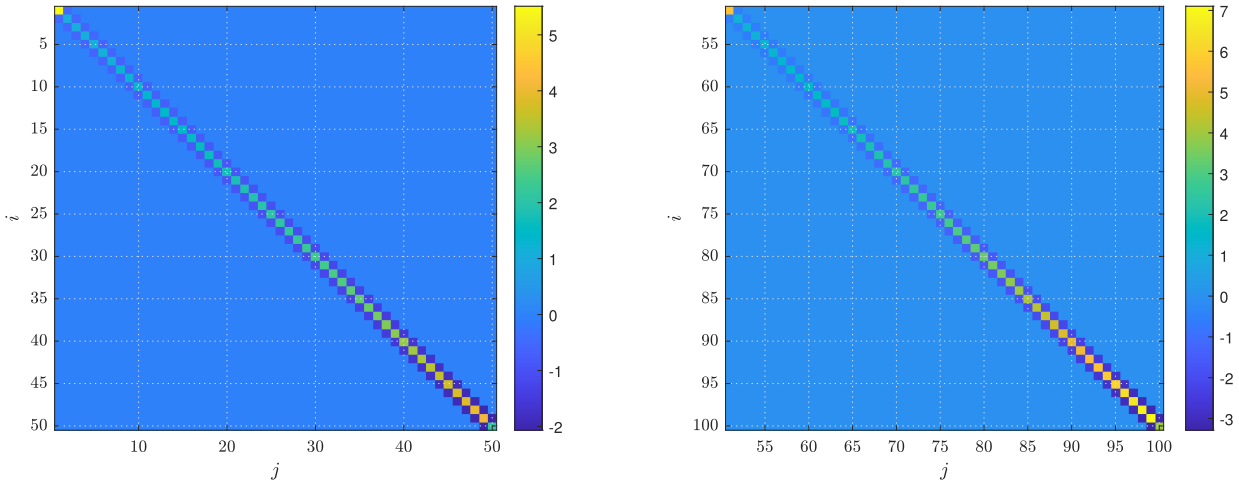



**FIGURE 2.** Input-output diabetic dataset used for identification.




(a) $\Lambda_{ij}\ \forall i < M_u, j < M_u$ corresponding to the insulin administration input (b) $\Lambda_{ij}\ \forall i \geq M_u, j \geq M_u$ corresponding to the carbohydrate intake input

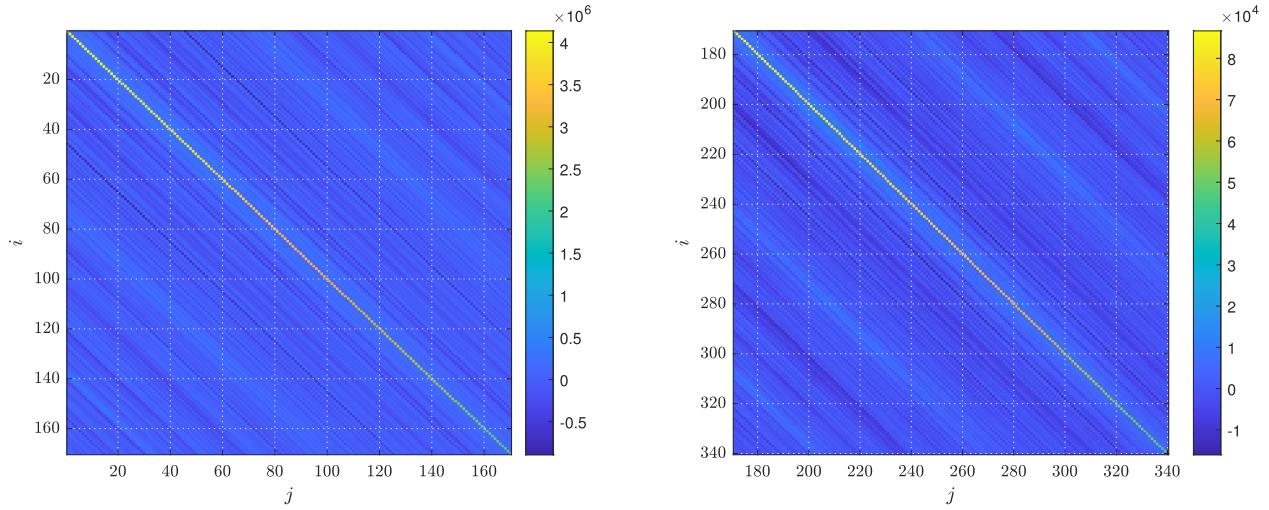
**FIGURE 3.** Entries of matrix $\Lambda$.

In the context of assessing the qualitative benefits of the generalized least squares method, the impulse response coefficients estimated assuming $\mathcal{Q}^{-1} = I$ are plotted in Fig. 7 to demonstrate the deterioration of estimate quality if the ordinary least squares method is applied. It can be concluded that the estimated impulse responses were less organized and showed poor compliance with the physiology despite applied regularization.

The next comparison is focused on the importance of regularization as one can observe significant qualitative differences between the regularized estimate in Fig. 5 and the non-regularized estimate in Fig. 8. In detail, the smoothness of the non-regularized estimate was not satisfactory, so the impulse responses did not have a realistic shape that would correspond to the physiology-based expectations. Due to the lack of regularization and the consequent increased variances of the estimate, the confidence intervals were significantly wider.

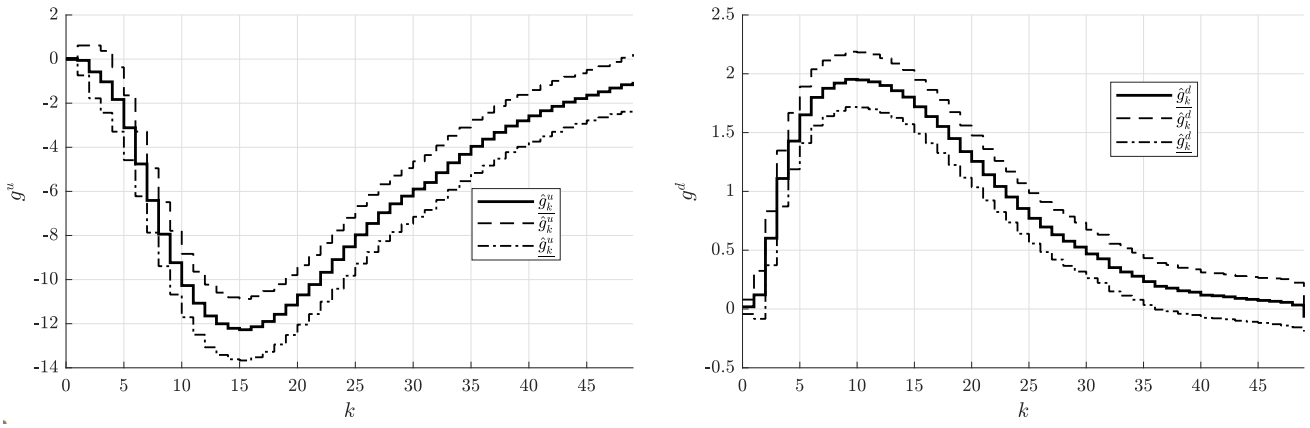
The orders of the empirical parametric model (56) were chosen as $n_{B^u} = 4, n_{B^d} = 3, n_{A^u} = 4, n_{A^d} = 3$ for the parametric approximation (59) and its solution (62). The obtained results are documented in Fig. 9 in a graphical form via plotting the impulse responses of the parametric model (56) and the source impulse responses from Fig. 5.

We will continue the comparisons by identifying the parameters of the two-input ARX model according to the straightforward prediction error approach based on the ordinary least squares method. For more details on this conservative




(a) $\left(\mathcal{Q}^{-1}\right)_{ij}\ \forall i < P, j < P$ corresponding to the insulin administration input (b) $\left(\mathcal{Q}^{-1}\right)_{ij}\ \forall i \geq P, j \geq P$ corresponding to the carbohydrate intake input

**FIGURE 4.** Elements of matrix $\mathcal{Q}^{-1}$.



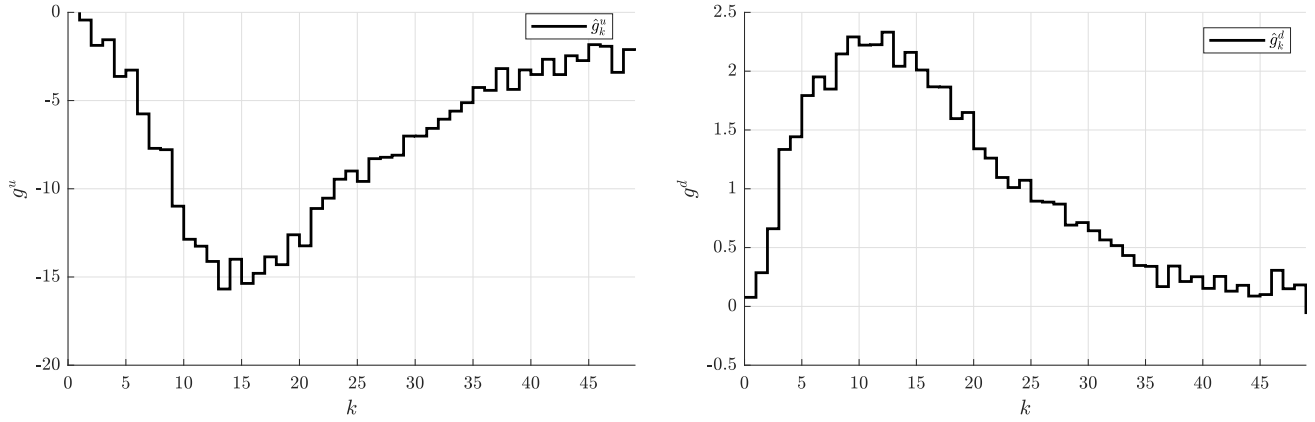

(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

**FIGURE 5.** Estimated impulse response coefficients of the insulin administration effect submodel $\hat{g}^u$ and the carbohydrate intake effect submodel $\hat{g}^d$ together with the corresponding confidence interval $\bar{\hat{g}}, \underline{\hat{g}}$.

approach, we refer the interested reader to [7]. Note that the ARX model can be considered a special case of the parametric model (56) under the assumption of common denominator $A^u(z) = A^d(z) = A(z)$ for both submodels. The orders of the ARX model were chosen as $n_{B^u} = 4$, $n_{B^d} = 3$, $n_A = 4$ for this experiment. The impulse responses of the identified ARX model are plotted in Fig. 9. Although the presented impulse responses appear to be smooth and relatively compliant with the physiology at first sight, the main problem is that both suggest for the presence of same time constants and the autoregressive dynamics that cause the peak time of insulin administration effect and carbohydrate intake effect to be almost identical. This experience is in conflict with the physiology and with the previously presented results.

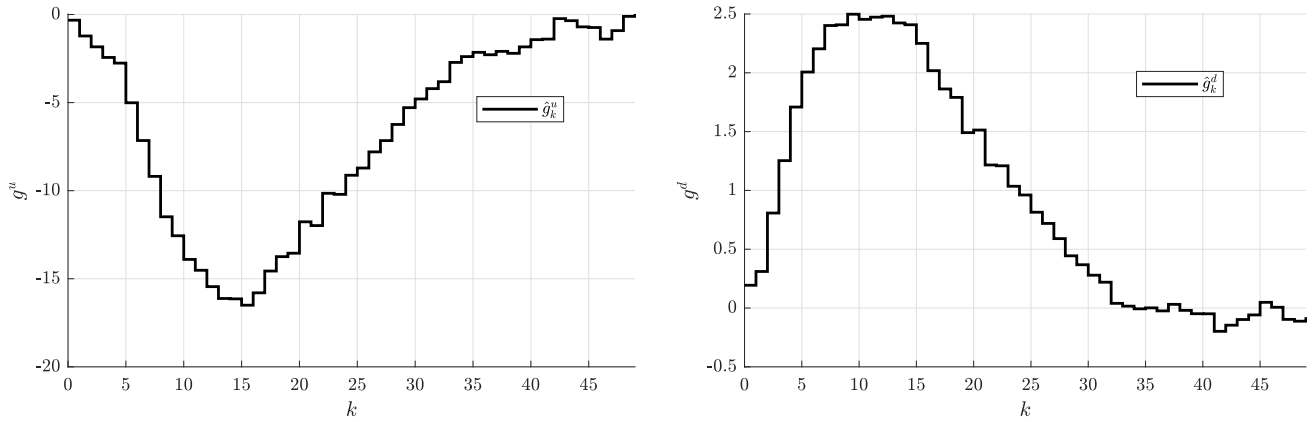
The comparison will be concluded by performing the identification of transfer function-based empirical model pursuing the predictive identification strategy proposed in [22]. We recall that the referred strategy was based on numeric minimization of the model multiple step-ahead prediction error formulated as a constrained nonlinear least squares problem. The impulse responses of thus identified model can be seen in Fig. 11 where one can notice that contrary to the impulse responses of the ARX model presented in Fig. 10, the structure with independent denominators is capable of accurately modelling the systems with significantly different time constants of both submodels.

The experiment will be concluded by validating the chosen identified models in a direct comparison of the predicted



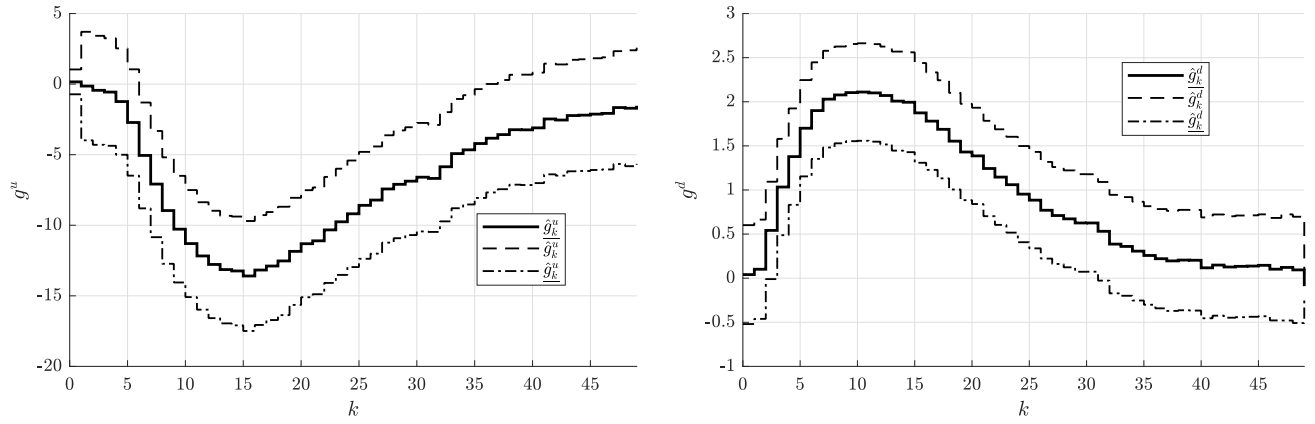

(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

**FIGURE 6.** Estimated impulse response coefficients of the insulin administration effect submodel $\hat{g}^u$ and the carbohydrate intake effect submodel $\hat{g}^d$ obtained by solving the regular system of linear equations as $\hat{g} = \hat{\mathcal{R}}_{\mathcal{U}}^{-1} \hat{\mathcal{R}}_{\mathcal{Y}}$.




(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

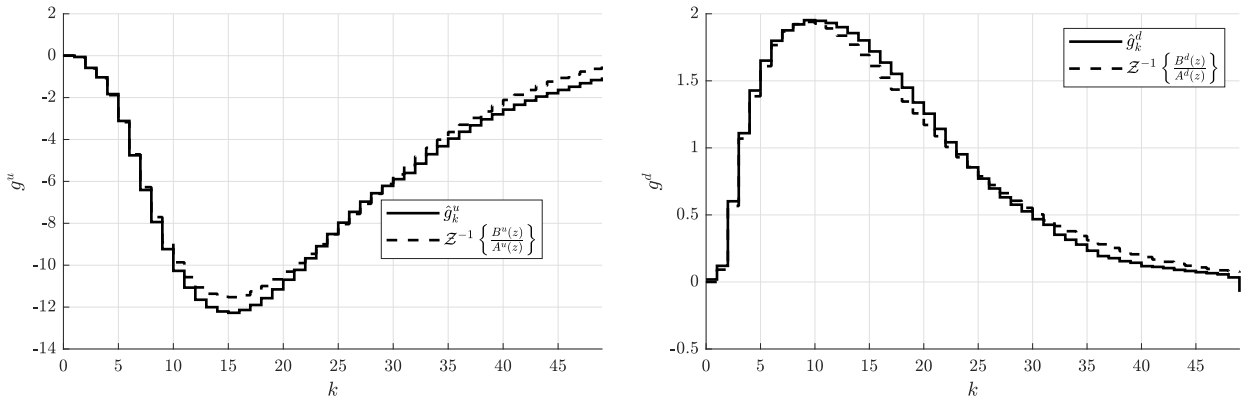
**FIGURE 7.** Estimated impulse response coefficients of the insulin administration effect submodel $\hat{g}^u$ and the carbohydrate intake effect submodel $\hat{g}^d$ under the assumption of unitary variance $\mathcal{Q}^{-1} = I$.




(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

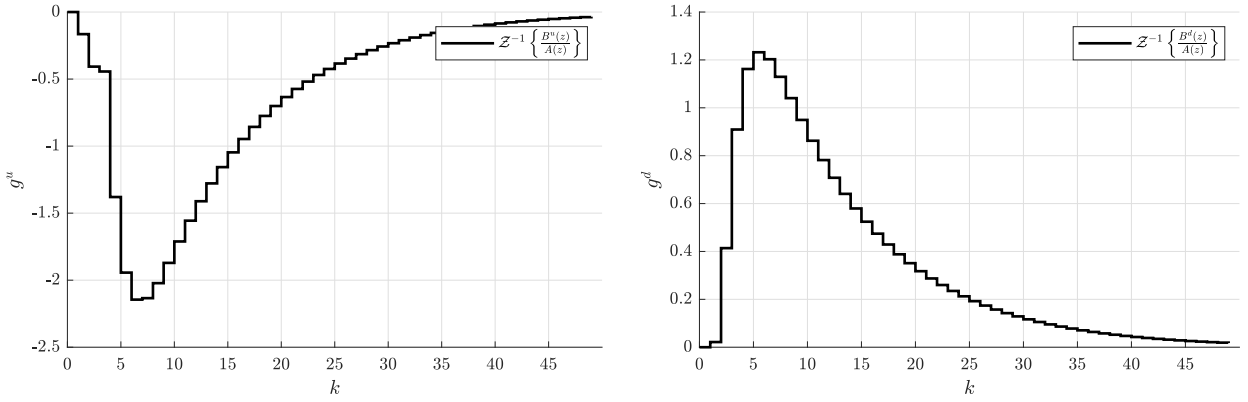
**FIGURE 8.** Estimated impulse response coefficients of the insulin administration effect submodel $\hat{g}^u$ and the carbohydrate intake effect submodel $\hat{g}^d$ together with the corresponding confidence interval $\overline{\hat{g}}$, $\underline{\hat{g}}$ under the assumption of no regularization $\Lambda = 0$.

output to the virtually measured glycemia. These long-term predictions will be performed for the nonparametric model (10) with the impulse responses from Fig. 5 and also for its parametric approximation with the corresponding impulse



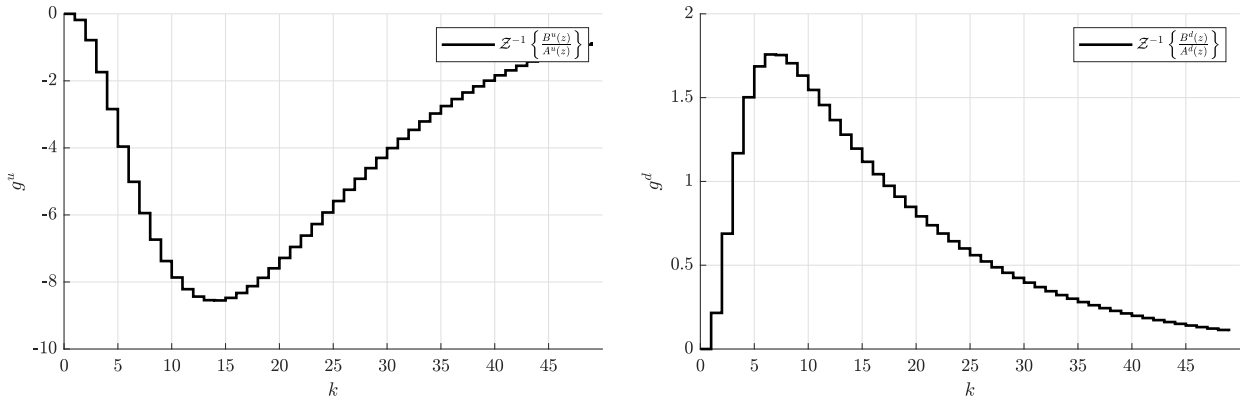

(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

**FIGURE 9.** Comparison of the estimated impulse response coefficients of the insulin administration effect submodel $\hat{g}^u$ and the carbohydrate intake effect submodel $\hat{g}^d$ to the impulse responses $\mathcal{Z}^{-1}\left\{\frac{B^u(z)}{A^u(z)}\right\}$, $\mathcal{Z}^{-1}\left\{\frac{B^d(z)}{A^d(z)}\right\}$ of the approximate parametric model (56).




(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

**FIGURE 10.** Impulse response of the insulin administration effect submodel $\mathcal{Z}^{-1}\left\{\frac{B^u(z)}{A(z)}\right\}$ and the carbohydrate intake effect submodel $\mathcal{Z}^{-1}\left\{\frac{B^d(z)}{A(z)}\right\}$ of the estimated two-input ARX model.




(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

**FIGURE 11.** Impulse response of the insulin administration effect submodel $\mathcal{Z}^{-1}\left\{\frac{B^u(z)}{A^u(z)}\right\}$ and the carbohydrate intake effect submodel $\mathcal{Z}^{-1}\left\{\frac{B^d(z)}{A^d(z)}\right\}$ estimated using the predictive identification approach adopted from [22].

responses from Fig. 9. For both models, the scenario representing the predictions based on the validation dataset, which is different from the identification dataset from Fig. 2, can be seen in Fig. 12. Finally, the performance of the identified

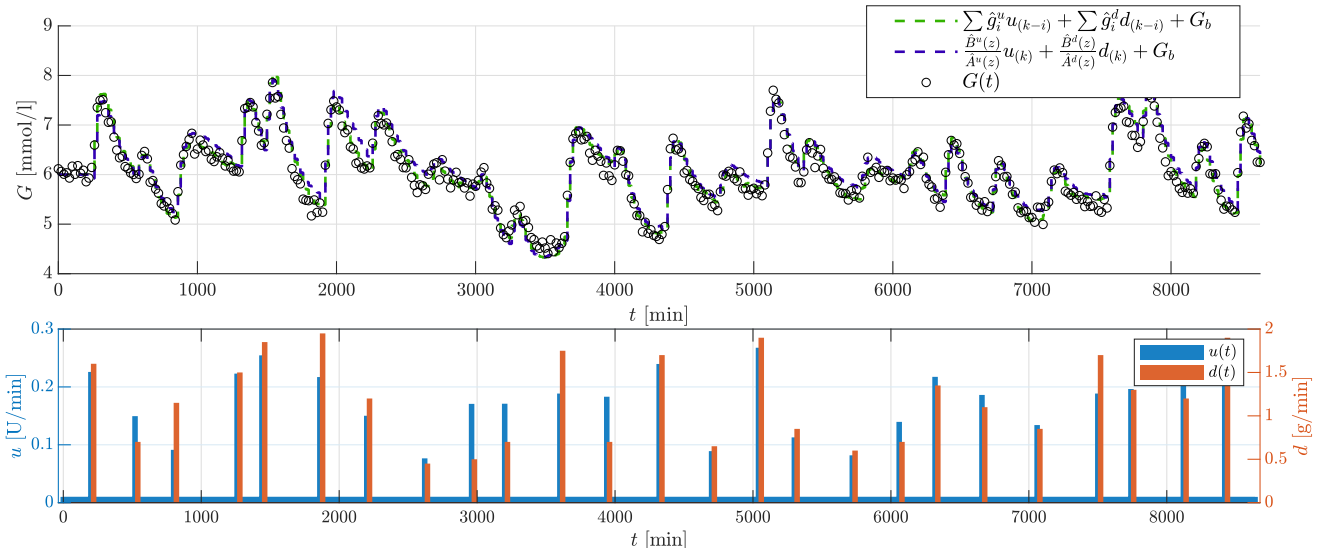



**FIGURE 12.** Prediction of the glycemia response for the validation dataset using the nonparametric model and the approximate parametric model.

**TABLE 1.** Performance metric of the identified models evaluated for the identification and the validation dataset.

| $Q$ | nonparametric model | parametric model |
|---|---|---|
| identification dataset | $1.6951 \times 10^{-2}$ | $1.9783 \times 10^{-2}$ |
| validation dataset | $2.2115 \times 10^{-2}$ | $3.7928 \times 10^{-2}$ |

models can be quantified by the metric

$$Q = \frac{1}{N} \sum_{k=1}^{N} \left[ \hat{y}_{(k)} - y_{(k)} \right]^2 , \tag{68}$$

where $\hat{y}_{(k)}$ is the predicted model output and $y_{(k)}$ is the measured output. Evaluating the metric (68) for the identified nonparametric model and the approximated parametric model yields the results summarized in Table 1. The results demonstrated that both models are capable to accurately predict glycemia in the long-term sense, while the parametric model performed slightly worse in both scenarios.

## VIII. CONCLUSION

This work presented a correlation-based identification algorithm to estimate impulse response coefficients of the two-input linear nonparametric empirical model of type 1 diabetes. This algorithm deals with the deconvolution problem and represents an alternative to the traditional identification techniques based on the least squares minimization of the model single step-ahead prediction error.

The most significant original contributions of this paper include the derivation of the generalized form of the Wiener-Hopf equation for the continuous-time model with two inputs. Based on the discrete-time equivalent of the Wiener-Hopf equation, the paper featured the linear equation system that resulted in formulation of the equivalent regression system. By solving this regression system in the least squares sense, the coefficients of impulse responses were estimated. In contrast to the state-of-the-art correlation-based identification methods available in the literature, we rigorously investigated the statistical properties of the correlation function estimate, concluding that it needs to be treated as uncertain and correlated in the formulation of the identification problem in order to ensure the optimal performance. The generalized least squares method was employed to obtain the parameter estimate with minimal variance instead of using the ordinary least squares method, which can typically be found in the literature. We also highlight that the covariance matrix of the random vector in the regression system was derived, as this matrix was vital to evaluate the generalized least squares method. As an extension to the proposed method, we designed an auxiliary algorithm for approximating the estimated impulse responses by the parametric transfer function model, again conveniently using the generalized least squares method while exploiting the covariance matrix of estimated impulse responses. The combination of the initial nonparametric model identification with the subsequent parametric model approximation can be seen as a full-featured two-step identification procedure for multiple-input transfer function-based models that have different feedback dynamics for each of the inputs. We stress that the aforementioned combination is especially suitable for identifying empirical models of type 1 diabetes due to its ability to capture distinct dynamics of the insulin administration effect and the carbohydrate intake effect, but also because of the existence of analytical solutions for both estimation steps, what significantly simplifies the possible real-time implementation. Moreover, three types of regularization were applied to obtain a smoother impulse response and also to provide causality and asymptotic stability of the identified model.

The results of the simulation-based experiment confirmed that a properly tuned identification algorithm can yield satisfactory and valid estimates of impulse functions even under demanding conditions caused by noisy glycemia measurements and highly correlated input signals typical for conventional insulin therapy.

By studying the obtained inverse $\mathcal{Q}^{-1}$ of the covariance matrix, we concluded that it had a quasi-diagonal structure, while the diagonal entries were decreasing with an increasing lag number. This structure participated in significant weighting of residuals in the generalized least squares method, but the decorrelating effect was not so prominent.

The presented comparison between the regularized and the non-regularized estimates of the impulse functions clearly indicated the qualitative superiority of the proposed combined triple regularization over the non-regularized estimate. In particular, the estimated impulse responses were smoother and appeared to be more organized if regularization was applied, making the estimate more valid and physiology-compliant. Concerning the character of the identified impulse responses, they were both stable and aperiodic, and each had a uniform sign that was compatible with the basic physiology. There was also observed a delayed and longer-lasting action of insulin administration compared to a faster and shorter-lasting effect of carbohydrate intake. We also demonstrated the importance of using the generalized least squares method, which was compared to the ordinary least squares method, concluding that the proposed strategy of minimizing the estimate variance outperforms the traditional approach. Moreover, the identified nonparametric and the approximated parametric models were both capable of accurately predicting glycemia over a long time period. As expected, the approximated parametric model exhibited slightly worse performance.

The key findings presented in this paper can be found useful within further studies focused on the predictive control-based artificial pancreas in patients with type 1 diabetes where model individualization is the essential prerequisite to achieve high control performance. Possible future improvements of this work may involve adding more sophisticated regularization kernels or even robustyfying the identification algorithm by penalizing deviations of the estimate from the expected mean-population impulse responses. It is also potentially feasible to design a recursive version of the proposed identification algorithm.

## REFERENCES

[1] M. Dodek and E. Miklovicova, "Optimal state estimation for the artificial pancreas," in *Proc. 23rd Int. Carpathian Control Conf. (ICCC)*, May 2022, pp. 88–93.

[2] J. O. Orozco-López, C. E. Castañeda, A. Rodríguez-Herrero, G. García-Sáez, and E. Hernando, "Linear time-varying Luenberger observer applied to diabetes," *IEEE Access*, vol. 6, pp. 23612–23625, 2018.

[3] M. Dodek and E. Miklovicová, "Maximizing performance of linear model predictive control of glycemia for T1DM subjects," *Arch. Control Sci.*, vol. 32, no. 2, pp. 305–333, 2022.

[4] L. Kovacs, G. Eigner, M. Siket, and L. Barkai, "Control of diabetes mellitus by advanced robust control solution," *IEEE Access*, vol. 7, pp. 125609–125622, 2019.

[5] G. De Nicolao, L. Magni, C. D. Man, and C. Cobelli, "Modeling and control of diabetes: Towards the artificial pancreas," *IFAC Proc. Volumes*, vol. 44, no. 1, pp. 7092–7101, Jan. 2011.

[6] R. Sánchez-Peña and D. Chernavvsky, *Artificial Pancreas: Current Situation and Future Directions*. New York, NY, USA: Academic, Apr. 2019.

[7] L. Ljung, *System Identification: Theory for the User* (Prentice-Hall information and system sciences series). Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[8] D. A. Finan, H. Zisser, L. Jovanovic, W. C. Bevier, and D. E. Seborg, "Identification of linear dynamic models for type 1 diabetes: A simulation study," *IFAC Proc. Volumes*, vol. 39, no. 2, pp. 503–508, 2006.

[9] D. A. Finan, C. C. Palerm, F. J. Doyle, D. E. Seborg, H. Zisser, W. C. Bevier, and L. Jovanovic, "Effect of input excitation on the quality of empirical dynamic models for type 1 diabetes," *AIChE J.*, vol. 55, no. 5, pp. 1135–1146, May 2009.

[10] M. Rebro, M. Tárník, and J. Murgaš, "Glycemia prediction accuracy of simple linear models with online parameter identification," *Int. Rev. Model. Simulations*, vol. 9, no. 5, p. 367, Oct. 2016.

[11] D. A. Finan, C. C. Palerm, F. J. Doyle, H. Zisser, L. Jovanovic, W. C. Bevier, and D. E. Seborg, "Identification of empirical dynamic models from type 1 diabetes subject data," in *Proc. Amer. Control Conf.*, Jun. 2008, pp. 2099–2104.

[12] H. Kirchsteiger, S. Polzer, R. Johansson, E. Renard, and L. Del Re, "Direct continuous time system identification of MISO transfer function models applied to type 1 diabetes," in *Proc. IEEE Conf. Decis. Control Eur. Control Conf.*, Dec. 2011, pp. 5176–5181.

[13] H. Kirchsteiger, R. Johansson, E. Renard, and L. D. Re, "Continuous-time interval model identification of blood glucose dynamics for type 1 diabetes," *Int. J. Control*, vol. 87, no. 7, pp. 1454–1466, Jul. 2014.

[14] H. Kirchsteiger, G. C. Estrada, S. Pölzer, E. Renard, and L. Del Re, "Estimating interval process models for type 1 diabetes for robust control design," *IFAC Proc. Volumes*, vol. 44, no. 1, pp. 11761–11766, 2011.

[15] M. Tárník, V. Bátora, J. B. Jørgensen, D. Boiroux, E. Miklovičová, T. Ludwig, I. Ottinger, and J. Murgaš, "Remarks on models for estimating the carbohydrate to insulin ratio and insulin sensitivity in T1DM," in *Proc. Eur. Control Conf. (ECC)*, Jul. 2015, pp. 31–36.

[16] C. Toffanin, S. Del Favero, E. M. Aiello, M. Messori, C. Cobelli, and L. Magni, "MPC model individualization in free-living conditions: A proof-of-concept case study," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1181–1186, Jul. 2017.

[17] C. Toffanin, S. Del Favero, E. M. Aiello, M. Messori, C. Cobelli, and L. Magni, "Glucose-insulin model identified in free-living conditions for hypoglycaemia prevention," *J. Process Control*, vol. 64, pp. 27–36, Apr. 2018.

[18] C. Toffanin, E. M. Aiello, S. Del Favero, C. Cobelli, and L. Magni, "Multiple models for artificial pancreas predictions identified from free-living condition data: A proof of concept study," *J. Process Control*, vol. 77, pp. 29–37, May 2019.

[19] C. Toffanin, E. M. Aiello, C. Cobelli, and L. Magni, "Hypoglycemia prevention via personalized glucose-insulin models identified in free-living conditions," *J. Diabetes Sci. Technol.*, vol. 13, no. 6, pp. 1008–1016, Nov. 2019.

[20] R. Hovorka, V. Canonico, L. J. Chassin, U. Haueter, M. Massi-Benedetti, M. O. Federici, T. R. Pieber, H. C. Schaller, L. Schaupp, T. Vering, and W. M. Wilinska, "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes," *Physiolog. Meas.*, vol. 25, no. 4, pp. 905–920, Jul. 2004.

[21] I. Aljamaan and I. Al-Naib, "Prediction of blood glucose level using nonlinear system identification approach," *IEEE Access*, vol. 10, pp. 1936–1945, 2022.

[22] M. Dodek and E. Miklovičová, "Physiology-compliant empirical model for glycemia prediction," *Int. Rev. Autom. Control*, vol. 14, no. 6, p. 310, Nov. 2021.

[23] S. Del Favero, G. Pillonetto, C. Cobelli, and G. De Nicolao, "A novel nonparametric approach for the identification of the glucose-insulin system in type 1 diabetic patients," *IFAC Proc. Volumes*, vol. 44, no. 1, pp. 8340–8346, Jan. 2011.

[24] M. Messori, C. Toffanin, S. Del Favero, G. Nicolao, C. Cobelli, and L. Magni, "A nonparametric approach for model individualization in an artificial pancreas," *IFAC-PapersOnLine*, vol. 48, pp. 225–230, Dec. 2015.

[25] M. Messori, C. Toffanin, S. Del Favero, G. De Nicolao, C. Cobelli, and L. Magni, "Model individualization for artificial pancreas," *Comput. Methods Programs Biomed.*, vol. 171, pp. 133–140, Apr. 2019.
[26] R. S. Parker, F. J. Doyle, and N. A. Peppas, "A model-based algorithm for blood glucose control in type I diabetic patients," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 148–157, Feb. 1999.
[27] C. Fabris and B. Kovatchev, *Glucose Monitoring Devices: Measuring Blood Glucose to Manage and Control Diabetes*. Amsterdam, The Netherlands: Elsevier Science, 2020.
[28] H. Kirchsteiger, J. Jørgensen, E. Renard, and L. Del Re, *Prediction Methods for Blood Glucose Concentration: Design, Use and Evaluation* (Lecture Notes in Bioengineering). Cham, Switzerland: Springer, 2016.
[29] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge, U.K.: Cambridge Univ. Press, 2006.
[30] D. Wang and C. D. Johnson, "A comparison study of some impulse-response identification methods," in *Proc. 21st Southeastern Symp. Syst. Theory*, 1989, pp. 52–56.
[31] G. Jenkins and D. Watts, *Spectral Analysis and Its Applications Time Series Analysis and Digital Signal Processing*. Toronto, ON, USA: Holden-Day, 1969.
[32] H. Pishro-Nik, *Introduction to Probability, Statistics, and Random Processes*. Athens, Greece: Kappa Research, 2014. [Online]. Available: https://www.probabilitycourse.com
[33] G. Box, G. Jenkins, G. Reinsel, and G. Ljung, *Time Series Analysis: Forecasting and Control Probability and Statistics*. Hoboken, NJ, USA: Wiley, 2015.
[34] T. Amemiya, "Generalized least squares theory," *Advanced Econometrics*. Cambridge, U.K.: Cambridge Univ. Press, 1985, pp. 181–185.
[35] W. Ling and D. Rivera, "Multivariable impulse response estimation via correlation analysis and its application to automated system identification," in *Proc. 11th IFAC Symp. Syst. Identificat. (SYSID)*, vols. 1–3, Y. Sawaragi and S. Sagara, Eds. Kitakyushu, Japan, 1998, pp. 1399–1404.
[36] Z. Wang, Q. Jin, and X. Liu, "Iteratively reweighted correlation analysis method for robust parameter identification of multiple-input multiple-output discrete-time systems," *IET Signal Process.*, vol. 10, no. 5, pp. 549–556, Jul. 2016.
[37] A. Marconato, M. Schoukens, and J. Schoukens, "Filter-based regularisation for impulse response modelling," *IET Control Theory Appl.*, vol. 11, no. 2, pp. 194–204, Jan. 2016.
[38] A. Chiuso, T. Chen, L. Ljung, and G. Pillonetto, "Regularization strategies for nonparametric system identification," in *Proc. 52nd IEEE Conf. Decis. Control*, Dec. 2013, pp. 6013–6018.
[39] T. Chen, H. Ohlsson, and L. Ljung, "On the estimation of transfer functions, regularizations and Gaussian processes-revisited," *Automatica*, vol. 48, no. 8, pp. 1525–1535, 2012.
[40] M. Cescon and R. Johansson, "Linear modeling and prediction in diabetes physiology," in *Data-driven Modeling for Diabetes* Lecture Notes in Bioengineering, V. Marmarelis and G. Mitsis, Ed. Berlin, Germany: Springer, 2014, pp. 187–222.
[41] S. Schmidt and K. Nørgaard, "Bolus calculators," *J. Diabetes Sci. Technol.*, vol. 8, no. 5, pp. 1035–1041, Sep. 2014.
[42] N. Cankaya and O. Aydogdu, "Three parameter control algorithm for obtaining ideal postprandial blood glucose in type 1 diabetes mellitus," *IEEE Access*, vol. 8, pp. 152305–152315, 2020.
[43] C. K. Boughton, S. Hartnell, J. M. Allen, and R. Hovorka, "The importance of prandial insulin bolus timing with hybrid closed-loop systems," *Diabetic Med.*, vol. 36, no. 12, pp. 1716–1717, Dec. 2019.
[44] C. Dalla Man, R. A. Rizza, and C. Cobelli, "Meal simulation model of the glucose-insulin system," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 10, pp. 1740–1749, Sep. 2007.
[45] C. D. Man, R. A. Rizza, and C. Cobelli, "Mixed meal simulation model of glucose-insulin system," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 307–310.
[46] L. Magni, D. M. Raimondo, L. Bossi, C. Dalla Man, G. De Nicolao, B. Kovatchev, and C. Cobelli, "Model predictive control of type 1 diabetes: An in silico trial," *J. Diabetes Sci. Technol.*, vol. 1, no. 6, pp. 804–812, Nov. 2007.
[47] C. Cobelli, C. Dalla Man, G. Sparacino, L. Magni, G. De Nicolao, and B. P. Kovatchev, "Diabetes: Models, signals, and control," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 54–96, 2009.

**MARTIN DODEK** was born in 1995. He received the M.Sc. degree in robotics and cybernetics from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, in 2020. He is currently pursuing the Ph.D. degree. His research interests include predictive control, system identification, optimization, and biocybernetics of type 1 diabetes mellitus.

**EVA MIKLOVIČOVÁ** received the M.Sc. and Ph.D. degrees in automation from the Slovak University of Technology in Bratislava, in 1990 and 1997, respectively. Currently, she is with the Faculty of Electrical Engineering and Information Technology, Institute of Robotics and Cybernetics, STU in Bratislava. Her research interests include predictive control, adaptive control, system modeling, and identification.

**MARIÁN TÁRNÍK** was born in Levice, Slovakia, in 1986. He received the Bachelor of Science degree from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, in 2008, and the Master of Science degree in technical cybernetics and the Ph.D. degree from the Slovak University of Technology in Bratislava, in 2010 and 2013, respectively. Since 2013, he has been as a Researcher. His research interests include adaptive control, Lyapunov design approach, and modeling and control of type one diabetes mellitus.