IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 13 February 2024, accepted 25 February 2024, date of publication 4 March 2024, date of current version 12 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3373068

RESEARCH ARTICLE

Robust Online Correlation Method for Identification of a Nonparametric Model of Type 1 Diabetes

MARTIN DODEK AND EVA MIKLOVIČOVÁ
Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, 811 07 Bratislava, Slovakia

Corresponding author: Martin Dodek (martin.dodek@stuba.sk)

This work was supported in part by the call for doctoral students and young researchers of Slovak University of Technology in Bratislava to start a research career (Impulsive Control of Biosystems) under Grant 23-03-01-B; and in part by the call for doctoral students and young researchers, funded by Plán obnovy a odolnosti Slovenskej republiky (POO SR) under Grant 09I03-03-V05.

ABSTRACT The paper presents an online version of the identification method for estimating the impulse responses in the case of a two-input single-output linear empirical model of type 1 diabetes that allows us to adapt the model parameters due to the intra-subject time variability in real time. The method builds on and augments our original research by providing important enhancements concerning the online parameter estimation, recursive formulation of essential equations, improved regularization, and new effective approaches to numerically solve the estimation problem. Recursive equations are derived to update the covariance matrix of the sample cross-correlation function, as well as the inverse of this covariance matrix, where the customized Sherman-Morrison formula was considered. To efficiently update the parameter estimate at each sample while avoiding direct calculation of the Hessian matrix inverse, two alternative strategies are proposed to be applied instead. The first is based on the numeric minimization by the conjugate gradient method, whereas the second takes advantage of the Schulz method to approximate the inverse Hessian matrix. As a result, all steps of the identification algorithm were designed so that only basic linear operations are required. Features to robustify the estimate were also involved, as the optimal regularization strategies based on the inverse of the covariance matrix of the actual parameter distribution and the inter-sample parameter drift were applied. In the end of the paper, a series of simulation-based experiments was carried out to assess the effectiveness of the proposed method and to demonstrate all of its aspects and important characteristics. The documented results showed that the method can yield valid estimates of impulse responses and also effectively adapt parameters in real time under the influence of time-varying physiology.

INDEX TERMS System identification, nonparametric model, correlation function, generalized least squares method, robust identification, online parameter estimate, conjugate gradient method, Schulz method, diabetes mellitus.

I. INTRODUCTION

Diabetes mellitus is a chronic metabolic disorder manifested by a persistently elevated blood glucose concentration, also called hyperglycemia, which causes numerous serious health issues for the subject. This problem is currently attracting the increasing attention of scientists from various fields, including control engineers. In this paper, we deal with type 1 diabetes, which can be characterized as an absolute insulin deficiency and is therefore considered the insulin-dependent form of this disorder. Since the diabetic subject is expected to be influenced by significant intra-subject time variability of the physiology-based characteristics, there is a need to design an online estimation algorithm that can

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

ensure necessary adaptation of the model parameters in real time.

It is known that the actual dynamics of glycemia is of an underlying stochastic nature, since the human body is subject to multiple random exogenous disturbances, which include the level of physical activity and stress factors, while significant roles are also played by the input uncertainties. The aforementioned stochastic effects use to be modelled by the corresponding process noise [1]. Therefore, it can be concluded that statistical and regression methods are suitable for parameter estimation.

An emerging significant practical reason for addressing the problem of identifying and personalizing empirical models of diabetes is to exploit this model within the model-based design of the optimal insulin bolus dosing policy and in the smart insulin bolus calculator algorithms. Moreover, it was also reported that the individualized model is vital for synthesis of model predictive control of glycemia [2], which leads to an implementation of the so-called artificial pancreas [3] and also for performing efficient state estimation [8].

Besides these sophisticated systematic applications, the estimated model in the form of impulse responses to insulin administration and carbohydrate intake can provide important insights into the dynamics of glycemia, which can be utilized in clinical practice in relatively straightforward and intuitive ways. One of the conclusions that can be drawn from these impulse functions is their magnitude, which is proportional to the static gain of the corresponding submodel. Then, these static gains can be used to determine the so-called insulin-carbohydrate ratio for the standard bolus calculator formula, and hence to make the therapeutic decisions by appropriately adjusting the strategy of insulin therapy. In addition to that, by analyzing the peak times of the estimated impulse responses, which typically result in the finding that the insulin administration effect is delayed compared to the carbohydrate intake effect, the timing of insulin administration can be adjusted accordingly to effectively compensate for the carbohydrate intake.

It is generally known that traditional identification approaches based on minimizing the single step-ahead prediction error of simple linear stochastic models in a combination with the ordinary least squares method or the recursive least squares method yield insufficient prediction performance, as well as poor physiology compliance and general validity. There obviously exist identification approaches that attempted to address the aforementioned issues [9], but these usually lack any closed-form analytical formula (solution) to obtain the parameter estimate, which results in increased computational complexity and problems with convergence of iterative schemes.

To address these challenges, the paper proposes a novel online identification method to obtain a time-varying model capable of long-term predicting glycemia under the influence of intra-subject variability while featuring the parameter estimate that can be determined by the means of the generalized least squares method. The theory builds on our study [10] that presented the generalization of the Wiener-Hopf equations for the two-input single-output systems and an offline correlation-based method to estimate the unknown impulse functions.

The paper has been divided into the following sections. In section II the preliminaries, the model structure, and the definition of exponentially weighted recursive estimate of the correlation function are presented. Section III introduces the generalized Wiener-Hopf equations and the formulation of the estimation problem based on the generalized least squares method, together with a detailed explanation of the regularization strategies. Section IV comprises the derivations of crucial recursive formulas for updating the covariance matrix of the correlation functions estimate, as well as the recursive relation for updating its inverse. In section V, the conjugate gradient optimization method is adopted to solve the identification problem numerically. Alternatively, section VI deals with the Schulz method to approximate the Hessian matrix inverse. Finally, section VII describes the experimental setup and discusses the results of simulation-based experiments designed to assess the actual effectiveness of all novel aspects of the proposed method applied to virtual diabetic data obtained under the influence of time-varying physiology.

The novel aspects presented in this work can be summarized as follows:

- Parameter estimate in terms of the generalized least squares method to minimize its variance
- Exponentially weighted sample cross-correlation function and its recursive version
- Recursive formulae for updating the covariance matrix of the sample cross-correlation function and its inverse
- New statistics-based optimal regularization strategies
- Effective numerical solution of the estimation problem by means of the conjugate gradient method

A. STATE OF THE ART

The design of methods and algorithms for estimating personalized mathematical models of glycemia dynamics in subjects with type 1 diabetes has been the subject of extensive scientific endeavor and research so far, as we will briefly analyze in this section.

The results considering basic linear stochastic models, such as ARX, ARMAX, and the Box-Jenkins model identified using standard algorithms minimizing the prediction error [11], were described in a study [12] and in an overview [13]. The authors claimed that the parameters of the ARMAX and the Box-Jenkins model had to be estimated using iterative methods, despite the fact that they are generally considered sensitive to convergence to a local minimum due to the non-convex nature of the optimization problem. Another important drawback is that the ARX model is not structurally compliant with basic physiology, since the autoregressive dynamics is shared between the insulin administration effect submodel and the carbohydrate intake effect submodel [10].

A comparison of stochastic linear models was presented in [14], yet in this case the authors applied online identification in terms of the recursive least squares method. Similarly, the adaptive ARMAX model was estimated at each time step by the recursive extended least squares method in [15] and by the weighted recursive least squares method in [16]. Another recursive identification technique for the ARX model by the means of normalized least mean squares method was presented in [17] and by the weighted recursive least squares method in [18].

However, as mentioned in [19], the single step-ahead prediction error criterion considered for the identification yields a model with poor validity and prediction performance.

Another problem of applying the recursive least squares method to adapt the parameters of linear models in real time is that the possible effects of exogenous random disturbances affecting the dynamics of glycemia [1] are hard to distinguish from the effects of time-varying system parameters. As a result, the online estimation algorithm will incorrectly adapt the parameters due to the effects of process noise.

Our recent work [9] proposed an offline identification strategy, which featured a numerical minimization of the multiple step-ahead prediction error of the model. It is important to mention that the compatibility of the model with the physiology of diabetes was ensured by performing the estimation in an appropriately constrained parameter space of the model poles, zeros, and gains.

From a different perspective, the papers [20] presented an offline identification method considering the continuous-time transfer function structure of the model. The aim of the identification was to find the parameter vector that minimizes the quadratic prediction error criterion by means of the Gauss-Newton algorithm. Although this approach appears to be interesting, there was a lack of online estimation features and the method was completely deterministic.

A similar continuous-time transfer function-based model was proposed in [23], where the system identification was performed pursuing an offline procedure based on the minimization of the sum of the squared prediction errors. However, since the residuals were a nonlinear function of the model parameters, the nonlinear least squares method formulation had to be considered, hence no closed analytical formula could be derived.

A nice review including the in-vivo results for methods of glycemia prediction in the context of machine learning and internet of medical things can be found in [27].

From online estimation approaches, in the older work [28] a complex nonlinear model was considered in combination with the Bayesian estimation applied to determine the time-varying model parameters. One of the important features was the online parameter reestimation to adapt the model due to the influence of intra-subject time-variability. In particular, the parameters reestimation was performed with each available sample while involving glucose measurements from the so-called ''learning window''. This strategy can be considered to be the most relevant to the goals specified in this paper, yet there was identified a physiology-based nonlinear model compared to a linear empirical nonparametric model considered in our work.

II. MODEL STRUCTURE AND PRELIMINARIES

In this paper, we will consider a two-input single-output linear nonparametric model with finite impulse responses. In the context of empirical modeling of glycemia dynamics in subjects with type 1 diabetes, the model output y [mmol/l] represents the deviation of glycemia from its steady-state value G_b [mmol/l]. The first input u [U/min] denotes the deviation of the insulin administration rate from the basal insulin dosing rate u_b [U/min], and the second input d [g/min] stands for the carbohydrate intake rate. The output of this model gets [10]

$$y_{(k)} = \sum_{i=0}^{M_u} g_i^u u_{(k-i)} + \sum_{i=0}^{M_d} g_i^d d_{(k-i)} + \epsilon_{(k)} \,, \tag{1}$$

where g_i^u are the impulse response coefficients of the insulin administration effect, g_i^d are the impulse response coefficients of the carbohydrate intake effect, and M_u, M_d are the assumed lengths of these impulse responses, respectively. Signal $\epsilon(t) \sim \mathcal{N}(0, \sigma_\epsilon^2)$ stands for the uncorrelated zero-mean random process, which is meant to represent continuous glucose monitoring sensor noise [29], [30] combined with the effects of other exogenous unmeasurable disturbances.

However, in contrast to the original offline algorithm [10], here we consider the parameter-varying structure of model (1). To this end, we will introduce the bracket notation $\cdot[\ ]$ for particular objects in order to disambiguate their instances in time and make the notation of crucial recursive relations neater throughout the paper. According to the outlined notation, $\cdot[k-1]$ refers to the previous sample and $\cdot[k]$ refers to the current sample with respect to time. The output of the parameter-varying model holds

$$y_{(k)} = \sum_{i=0}^{M_u} g_i^u[k] u_{(k-i)} + \sum_{i=0}^{M_d} g_i^d[k] d_{(k-i)} + \epsilon_{(k)} \,. \tag{2}$$

A. EXPONENTIALLY WEIGHTED ESTIMATE OF THE CROSS-CORRELATION FUNCTION

Before presenting the correlation-based identification method itself, the exponentially weighted sample cross-correlation function will be introduced as a necessary prerequisite.

The cross-correlation function $R_{xz}(n)$ of two general discrete-time infinite-length signals $x_{(k)}$, $z_{(k)}$ is, under the assumption of ergodicity, defined by the expectancy [31]

$$R_{xz}(n) = E\left\{x_{(k)} z_{(k-n)}\right\} \quad \forall k \in \mathbb{N} \,, \tag{3}$$

for the lag argument $n \in \mathbb{Z}$. If $x_{(k)} = z_{(k)}$, then $R_{xx}(n)$ is called the autocorrelation function. It is important to note that the

property

$$E\left\{x_{(k)}z_{(k+n)}\right\} = E\left\{x_{(l)}z_{(l+n)}\right\} \quad \forall k, l \in \mathbb{Z} \tag{4}$$

holds for (3) if $x_{(k)}$, $z_{(k)}$ are stationary and ergodic.

Therefore, if the underlying processes are not stationary, the ergodicity does not apply, and the cross-correlation function will be sample-dependent as

$$R_{xz}(n)[k] = E\left\{x_{(k)}z_{(k-n)}\right\} . \tag{5}$$

Now we introduce the exponentially weighted sample estimate of the cross-correlation function (3) obtained by processing N available samples while assuming the forgetting factor $0<\lambda<1$ as

$$\hat{R}_{xz}(n)[N] = \frac{1}{\sum_{i=1}^{N-n}\lambda^{(N-n-i)}}\sum_{i=1}^{N-n}\lambda^{(N-n-i)}x_{(i+n)}z_{(i)} , \tag{6}$$

where $n \in \mathbb{Z}$ is the integer lag argument satisfying $n < N$. Note that if $\lambda = 1$, then (6) is equivalent to the standard formula [32], which was considered in [10]. Formula (6) implies that the newest sample is weighted as $\lambda^{(N-n-N+n)} = \lambda^0 = 1$ while the oldest is, in the limit number of samples, weighted as $\lim_{N\to\infty}\lambda^{(N-n-1)}=0$ when estimating the cross-correlation function. This means that (6) is more suitable for estimating the sample-dependent cross-correlation function in the case of non-stationary processes.

However, in the case of a time-varying model (2), the estimate (6) cannot be claimed to be unbiased. In reality, obtaining an unbiased sample cross-correlation function would require to consider only the last term in the summation (6) as $\hat{R}_{xz}(n)[N] = x_{(N)}z_{(N-n)}$, which is totally impractical due to the very high variance of such an estimate. Therefore, formula (6) represents an estimate with a good bias-variance trade-off that can be affected by tuning the forgetting factor λ.

1) RECURSIVE FORM

To derive the recursive form of (6), which will be essential for online identification, we assume that the length of the processed timeseries increases with each sample as $N \leftarrow N+1$.

Then, the summation in (6) will be further denoted as

$$s(n)[N] = \sum_{i=1}^{N-n}\lambda^{(N-n-i)} . \tag{7}$$

More importantly, the recursive formula for effective updating of summation (7) can be derived as

$$s(n)[N] = \lambda s(n)[N-1] + 1 . \tag{8}$$

Finally, considering notation (7), the recursive relation to update the estimate (6) can be derived as

$$\hat{R}_{xz}(n)[N] = \frac{1}{s(n)[N]}\left(\lambda s(n)[N-1]\hat{R}_{xz}(n)[N-1] + x_{(N)}z_{(N-n)}\right). \tag{9}$$

It can be concluded that, due to the recursive formula (9), calculating the entire summation in (6) at each sample can be effectively avoided during online identification. The presented exponentially weighted recursive estimate of the cross-correlation function is one of the important novel features of this paper.

III. ESTIMATE OF THE IMPULSE RESPONSE COEFFICIENTS

To estimate the impulse response coefficients $g_u[N+1]$, $g_d[N+1]$ considering $N+1=k$ is the new sample and N is the previous sample, the cross-correlation functions of the system output with the inputs will be essential. The derivation of these cross-correlation functions results in a generalization of the Wiener-Hopf equation for a two-input single-output system [10].

Consider the cross-correlation function $R_{yu}(n)[N+1]$ of the time-varying model (2) as

$$R_{yu}(n)[N+1] = \sum_{i=0}^{M_u} g_i^u[N+1]R_{uu}(n-i)[N+1] + \sum_{i=0}^{M_d} g_i^d[N+1]R_{du}(n-i)[N+1] , \tag{10}$$

and the cross-correlation function $R_{yd}(n)[N+1]$, which can be derived as

$$R_{yd}(n)[N+1] = \sum_{i=0}^{M_u} g_i^u[N+1]R_{ud}(n-i)[N+1] + \sum_{i=0}^{M_d} g_i^d[N+1]R_{dd}(n-i)[N+1] . \tag{11}$$

Notice that $R_{yu}(n)[N+1]$ and $R_{yd}(n)[N+1]$ are dependent only on the current parameters $g^u[N+1]$, $g^d[N+1]$, which is quite convenient.

By replacing the true cross-correlation functions $R_{yu}(n)$, $R_{yd}(n)$ with their sample-based estimates $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$, we can claim that

$$\hat{R}_{yu}(n) = R_{yu}(n) + \zeta^{yu}(n) , \tag{12a}$$
$$\hat{R}_{yd}(n) = R_{yd}(n) + \zeta^{yd}(n) , \tag{12b}$$

which means that $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$ are random variables, while $\zeta^{yu}(n)$, $\zeta^{yd}(n)$ are their corresponding uncertainties.

Equations (10), (11) can form the equivalent linear regression system by considering the lag argument $n=0\ldots P$ [10] as

$$\begin{pmatrix}\hat{\mathcal{R}}_{yu}\\ \hat{\mathcal{R}}_{yd}\end{pmatrix} = \begin{pmatrix}\hat{\mathcal{R}}_{uu} & \hat{\mathcal{R}}_{du}\\ \hat{\mathcal{R}}_{ud} & \hat{\mathcal{R}}_{dd}\end{pmatrix}\begin{pmatrix}g^u\\ g^d\end{pmatrix} + \begin{pmatrix}\zeta^{yu}\\ \zeta^{yd}\end{pmatrix} , \tag{13}$$

where submatrices $\hat{\mathcal{R}}_{uu} \in \mathbb{R}^{P+1\times M_u+1}$, $\hat{\mathcal{R}}_{dd} \in \mathbb{R}^{P+1\times M_d+1}$, $\hat{\mathcal{R}}_{ud} \in \mathbb{R}^{P+1\times M_u+1}$, $\hat{\mathcal{R}}_{du} \in \mathbb{R}^{P+1\times M_d+1}$ have the general

structure

$$\hat{\mathcal{R}}_{xz} = \begin{pmatrix} \hat{R}_{xz}(0) & \hat{R}_{zx}(1) & \hat{R}_{zx}(2) & \dots & \hat{R}_{zx}(M_x) \\ \hat{R}_{xz}(1) & \hat{R}_{xz}(0) & \hat{R}_{zx}(1) & \dots & \hat{R}_{zx}(M_x-1) \\ \hat{R}_{xz}(2) & \hat{R}_{xz}(1) & \hat{R}_{xz}(0) & \dots & \hat{R}_{zx}(M_x-2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{xz}(P) & \hat{R}_{xz}(P-1) & \hat{R}_{xz}(P-2) & \dots & \hat{R}_{xz}(P-M_x) \end{pmatrix} , \tag{14}$$

and vectors $\hat{\mathcal{R}}_{yu} \in \mathbb{R}^{P+1\times 1}$, $\hat{\mathcal{R}}_{yd} \in \mathbb{R}^{P+1\times 1}$ get

$$\hat{\mathcal{R}}_{yu} = \begin{bmatrix} \hat{R}_{yu}(0) & \hat{R}_{yu}(1) & \dots & \hat{R}_{yu}(P) \end{bmatrix}^{\mathrm{T}} , \tag{15a}$$

$$\hat{\mathcal{R}}_{yd} = \begin{bmatrix} \hat{R}_{yd}(0) & \hat{R}_{yd}(1) & \dots & \hat{R}_{yd}(P) \end{bmatrix}^{\mathrm{T}} . \tag{15b}$$

We will further use the shorthand notation

$$\hat{\mathcal{R}}_{\mathcal{Y}} = \hat{\mathcal{R}}_{\mathcal{U}} g + \zeta \tag{16}$$

for (13).

It is important to note that the elements of regression matrix $\hat{\mathcal{R}}_{\mathcal{U}}$ and vector $\hat{\mathcal{R}}_{\mathcal{Y}}$ in the regression system (16) change completely at each iteration and can be effectively updated according to recursive formula (9).

The parameter vector $g \in \mathbb{R}^{M_u+M_d+2\times 1}$ can be formally defined as

$$g = \begin{bmatrix} g^u \\ g^d \end{bmatrix} , \tag{17}$$

where subvectors $g^u \in \mathbb{R}^{M_u+1\times 1}$ and $g^d \in \mathbb{R}^{M_d+1\times 1}$ get

$$g^u = \begin{bmatrix} g^u_0 & g^u_1 & g^u_2 & \dots & g^u_{M_u} \end{bmatrix}^{\mathrm{T}} , \tag{18a}$$

$$g^d = \begin{bmatrix} g^d_0 & g^d_1 & g^d_2 & \dots & g^d_{M_d} \end{bmatrix}^{\mathrm{T}} . \tag{18b}$$

As mentioned before, we will introduce the random vectors $\zeta^{yu} \in \mathbb{R}^{P+1\times 1}$, $\zeta^{yd} \in \mathbb{R}^{P+1\times 1}$ representing the errors (uncertainties) of the sample cross-correlation functions $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$ for $n=0 \dots P$ as

$$\begin{aligned} \zeta^{yu} &= \begin{bmatrix} \zeta^{yu}(0) & \zeta^{yu}(1) & \dots & \zeta^{yu}(P) \end{bmatrix}^{\mathrm{T}} \\ &= \begin{bmatrix} \hat{R}_{yu}(0) - R_{yu}(0) & \hat{R}_{yu}(1) - R_{yu}(1) \\ & \dots \hat{R}_{yu}(P) - R_{yu}(P) \end{bmatrix}^{\mathrm{T}} , \end{aligned} \tag{19}$$

$$\begin{aligned} \zeta^{yd} &= \begin{bmatrix} \zeta^{yd}(0) & \zeta^{yd}(1) & \dots & \zeta^{yd}(P) \end{bmatrix}^{\mathrm{T}} \\ &= \begin{bmatrix} \hat{R}_{yd}(0) - R_{yd}(0) & \hat{R}_{yd}(1) - R_{yd}(1) \\ & \dots \hat{R}_{yd}(P) - R_{yd}(P) \end{bmatrix}^{\mathrm{T}} . \end{aligned} \tag{20}$$

Vectors ζ^{yu}, ζ^{yd} can be joint into

$$\zeta = \begin{bmatrix} \zeta^{yu} \\ \zeta^{yd} \end{bmatrix} . \tag{21}$$

The maximal lag number P must satisfy the condition

$$M_u + M_d \; < 2P \ll 2N \; , \tag{22}$$

and should be chosen with regard to the character of estimated cross-correlation functions $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$.

A. STATISTICAL PROPERTIES OF THE PARAMETERS

In order to reflect the presence of inter-subject parametric variability typical for the diabetic patient population, we will assume that the actual parameter vector g follows the multivariate normal distribution. Therefore, the mean-population parameter vector $g_\mu \in \mathbb{R}^{M_u+M_d+2\times 1}$ will be defined as

$$g_\mu = E\{g\} \; . \tag{23}$$

The covariance matrix $\Psi \succ 0 \in \mathbb{R}^{M_u+M_d+2\times M_u+M_d+2}$, representing the statistical model of the inter-subject parametric variability, will be defined as

$$\Psi = \mathrm{cov}\,(g,g) = E\left\{ (g - E\{g\})\,(g - E\{g\})^{\mathrm{T}} \right\} , \tag{24}$$

and can be further divided into submatrices such that

$$\Psi = \begin{pmatrix} \Psi^{uu} & \Psi^{ud} \\ \Psi^{du} & \Psi^{dd} \end{pmatrix} . \tag{25}$$

In practice, the mean vector g_μ and the covariance matrix Ψ have to be estimated by studying a set of parameter estimates obtained experimentally from a sufficiently large group of subjects. Considering n_p different subjects and the corresponding estimated parameter vectors, where the p-th parameter vector of the dataset is indexed as $\hat{g}(p)$, then g_μ (23) can be estimated as the sample mean $\bar{g}$ [32], [33]

$$\bar{g} = \frac{1}{n_p} \sum_{p=1}^{n_p} \hat{g}(p) \, , \tag{26}$$

and the sample covariance matrix Ψ (24) can be obtained as [32]

$$\hat{\Psi} = \frac{1}{n_p - 1} \sum_{p=1}^{n_p} \left[\hat{g}\,(p) - \bar{g} \right] \left[\hat{g}\,(p) - \bar{g} \right]^{\mathrm{T}} . \tag{27}$$

Furthermore, in order to reflect the presence of intra-subject parametric variability, we will assume that the actual parameter vector is time-varying, so the bracket notation $g[i]$ will be used to distinguish its individual occurrences. Therefore, the actual parameter vector can be seen as a random process, while anticipating that its drift leads to no permanent bias, which can be noted as

$$E\{g[i]\} = g \; \forall i \, . \tag{28}$$

According to (28) the expectancy

$$E\{g[k] - g[k-1]\} = E\{g[k]\} - E\{g[k-1]\} = \mathbf{0} \, . \tag{29}$$

holds for the difference of two successive parameter vectors representing the inter-sample parameter change. The covariance matrix $\Phi \succ 0 \in \mathbb{R}^{M_u+M_d+2\times M_u+M_d+2}$ of the inter-sample parameter change $g[N+1] - g[N]$ is defined as

$$\begin{aligned} \Phi &= \mathrm{cov}\,(g[k] - g[k-1]) \\ &= E\left\{ (g[k] - g[k-1])\,(g[k] - g[k-1])^{\mathrm{T}} \right\} , \end{aligned} \tag{30}$$

and can be further divided into submatrices such that

$$\Phi = \begin{pmatrix} \Phi^{uu} & \Phi^{ud} \\ \Phi^{du} & \Phi^{dd} \end{pmatrix} . \tag{31}$$

To estimate this covariance matrix, formula

$$\hat{\Phi} = \frac{1}{n_p}\frac{1}{l-2}\sum_{p=1}^{n_p}\sum_{i=1}^{l-1}\left(\hat{g}(p)[i+1] - \hat{g}(p)[i]\right)\left(\hat{g}(p)[i+1] - \hat{g}(p)[i]\right)^{\mathrm{T}} \tag{32}$$

has to be applied to n_p sequences of the parameter estimates $\hat{g}(p)[i]$ with the length of l samples.

B. GENERALIZED LEAST SQUARES METHOD

To solve the linear regression system (16) and thus obtain the estimate of the parameter vector (17), the generalized least squares method [34] will be adopted. The rationale for choosing the generalized least squares method instead of the ordinary least squares method is that it provides the mechanisms of "weighting" and "decorrelating" of the residuals by the Cholesky decomposition of the inverse covariance matrix $\mathcal{Q}^{-1}$ of the noise vector ζ, thus minimizing the variance of the estimate [10]. Considering the generalized least squares method is also one of the novel aspects presented in this paper, compared to the standard identification approaches in the correlation framework.

The corresponding cost function customized by adding two regularization terms gets the quadratic form

$$J(\hat{g}) = \frac{1}{2}\Big[\left(\hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}}\hat{g}\right)^{\mathrm{T}}\mathcal{Q}^{-1}\left(\hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}}\hat{g}\right) + \left(\hat{g}-\bar{g}\right)^{\mathrm{T}}\alpha\Psi^{-1}\left(\hat{g}-\bar{g}\right) + \left(\hat{g}-\hat{g}[N]\right)^{\mathrm{T}}\beta\Phi^{-1}\left(\hat{g}-\hat{g}[N]\right)\Big], \tag{33}$$

where $\mathcal{Q} \succ 0 \in \mathbb{R}^{2(P+1)\times 2(P+1)}$ is the covariance matrix of the noise vector ζ, $\Psi^{-1} \succ 0$ (24), $\Phi^{-1} \succ 0$ (30) are positive definite regularization matrices, $\alpha \geq 0$, $\beta \geq 0$ are positive scalar weighting factors, $\bar{g}$ is the mean-population parameter vector (26), and $\hat{g}[N]$ is the parameter estimate from the last sample.

The gradient $\nabla_{\hat{g}} J(\hat{g}) = \mathbf{g}(\hat{g}) : \mathbb{R}^{M_u+M_d+2\times 1} \rightarrow \mathbb{R}^{M_u+M_d+2\times 1}$ of the cost function (33) with respect to the estimated parameter vector $\hat{g}$ can be derived as

$$\mathbf{g}(\hat{g}) = -\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\left(\hat{\mathcal{R}}_{\mathcal{Y}} - \hat{\mathcal{R}}_{\mathcal{U}}\hat{g}\right) + \alpha\Psi^{-1}\left(\hat{g}-\bar{g}\right) + \beta\Phi^{-1}\left(\hat{g}-\hat{g}[N]\right). \tag{34}$$

The Hessian matrix $\mathbf{H}$:$\mathbb{R}^{M_u+M_d+2\times 1} \rightarrow \mathbb{R}^{M_u+M_d+2\times M_u+M_d+2}$ gets

$$\mathbf{H} = \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}} + \alpha\Psi^{-1} + \beta\Phi^{-1}. \tag{35}$$

The Hessian matrix is independent of $\hat{g}$ (it is constant) and is positive definite since $\mathcal{Q}^{-1} \succ 0$, $\Psi^{-1} \succ 0$, $\Phi^{-1} \succ 0$, $\alpha > 0$, $\beta > 0$.

The optimal parameter estimate that minimizes cost function (33) can be obtained in the closed form based on the optimality condition $\mathbf{g}(\hat{g}) = \mathbf{0}$ as

$$\hat{g} = \mathbf{H}^{-1}\left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{Y}} + \alpha\Psi^{-1}\bar{g} + \beta\Phi^{-1}\hat{g}[N]\right). \tag{36}$$

One of the advantages of this correlation-based method over traditional prediction-error methods is that the computational complexity and dimensions of the estimation problem are not dependent on the number of processed samples N, and therefore the online estimation should be more feasible.

The recursive formula to effectively update matrix $\hat{\mathcal{R}}_{\mathcal{U}}$ and vector $\hat{\mathcal{R}}_{\mathcal{Y}}$ at each sample using the new output measurement and the input data can be derived based on (13), (15a), (15b) and the structure of partial submatrices (14), while applying the recursive formula (9) for the general cross-correlation/autocorrelation function estimate. However, to keep the paper fluent, this will be omitted.

Taking the expectancy operator to the estimate (36) yields

$$E\{\hat{g}\} = \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}} + \alpha\Psi^{-1} + \beta\Phi^{-1}\right)^{-1} \cdot \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}}g + \alpha\Psi^{-1}\bar{g} + \beta\Phi^{-1}\hat{g}[N]\right). \tag{37}$$

The above equation implies that the regularization terms in (33) induce some bias, since if $\alpha = \beta = 0$, then $E\{\hat{g}\} = g$.

The covariance matrix $\mathcal{P} \in \mathbb{R}^{M_u+M_d+2\times M_u+M_d+2}$ of estimate (36) can be derived as

$$\mathcal{P} = \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}} + \alpha\Psi^{-1} + \beta\Phi^{-1}\right)^{-1} \cdot \hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}}\left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}} + \alpha\Psi^{-1} + \beta\Phi^{-1}\right)^{-1}. \tag{38}$$

Due to the generalized least squares method, the estimate covariance matrix $\mathcal{P}$ can be considered minimal. It is important to note that if no regularization is applied, i.e. $\alpha = \beta = 0$, then the above formula reduces to

$$\mathcal{P} = \left(\hat{\mathcal{R}}_{\mathcal{U}}^{\mathrm{T}}\mathcal{Q}^{-1}\hat{\mathcal{R}}_{\mathcal{U}}\right)^{-1}, \tag{39}$$

which commonly occurs in the literature [34].

C. REGULARIZATION STRATEGIES

Regularization strategies use to be applied in order to involve some prior knowledge about the identified system in the estimation algorithm, decrease the variance of the estimate, prevent overfitting and improve the generalization of models by adding additional constraints or penalties to the optimization process. Regularization techniques integrated into the least squares method have a profound influence on parameter estimation robustness. By imposing additional constraints in the form of penalties, the susceptibility of parameter estimates to outliers or extreme values in the data is effectively mitigated. Regularization strategies not only foster more stable and generalized models, but also enhance the robustness of parameter estimation, fortifying the reliability of the inferred model coefficients [36].

Applying the regularization can be seen as penalizing certain properties that describe unlikely systems [37], and hence the robustness of the estimate should also be improved. As a drawback, regularization always induces some bias (see (37)), so a reasonable bias-variance trade-off must be chosen [38].

One way of involving prior knowledge in regularization is via the inverse Ψ^{-1} of the covariance matrix (24) of the actual parameter vector distribution, which can be considered the "optimal" regularization matrix [38]. Therefore the first regularization term $\left(\hat{g}-\bar{g}\right)^{\mathrm{T}}\alpha\Psi^{-1}\left(\hat{g}-\bar{g}\right)$ in cost function (33) penalizes the deviation of the parameter estimate $\hat{g}$ from the mean population value $\bar{g}$ (26).

Recall that the online parameter estimation was proposed primarily in order to reflect the time variability of the actual parameter vector, also referred to as intra-subject variability. Therefore, to robustify the identification algorithm and hence prevent unwanted excessive changes of the parameter estimate throughout the iterations due to the presence of outliers and deteriorated quality of input-output data, a dedicated regularization term $\left(\hat{g}-\hat{g}[N]\right)^{\mathrm{T}}\beta\Phi^{-1}\left(\hat{g}-\hat{g}[N]\right)$ is considered in cost function (33). This regularization strategy penalizes the inter-sample change $\hat{g}[N+1]-\hat{g}[N]$ of the estimate using the inverse Φ^{-1} of covariance matrix (30) of drift $g[N+1]-g[N]$ of the actual parameter vector.

Compared with the aforementioned optimal statistics-based regularization techniques, rather heuristic strategies, particularly the combination of regularizations to provide smoothness, stability, and causality, were originally proposed in [10].

D. TUNING THE ALGORITHM

When tuning the maximal lag P, the number of available samples N has to be considered since increasing the lag argument n in (6) truncates the summation interval to $N-n$ and thus might provide highly uncertain estimates of the correlation functions. Besides that, P also affects the complexity of the identification problem as the number of equations in (13) is $2P$. On the other hand, if $2P$ approaches M_u+M_d, the accuracy and validity of the estimated model deteriorates.

Concerning the lengths M_u, M_d of the estimated impulse responses, their choice should reflect the given sample time and the anticipated dynamics of the identified system such that the identified finite impulse response model sufficiently covers the decay phase of the impulse response.

However, the most crucial parameter to be tuned is the forgetting factor λ in the sample correlation function (6). By adjusting the forgetting factor λ, the adaptive properties of the identification algorithm can be affected such that a lower λ generally provides faster adaptation. Therefore, a lower λ should be applied in the case of highly time-varying systems, which require fast adaptation of parameters, while a higher λ is suitable for systems, the parameters of which are not expected to evolve significantly over time. It is worth noting that intensive forgetting leads to inconsistency and deteriorated accuracy of the model due to the phenomenon of discarding the valuable information in the older samples. Therefore, a reasonable trade-off between the model accuracy and the adaptation rate must be chosen carefully.

To adjust the strength of both regularization terms, the scalar parameters $\alpha > 0$ and $\beta > 0$ can be tuned. The tuning procedure can be carried out as follows: First, a nonregularized estimate of impulse responses is obtained by letting $\alpha = 0$, $\beta = 0$, while its quality and validity are visually assessed to roughly match the expected physiology-compliant responses. Then, α can be gradually increased until satisfactory shapes of the impulse responses are achieved. After doing so, α is fixed, and β can be gradually increased until acceptably smooth transitions between consecutive estimates of the impulse responses are obtained.

IV. COVARIANCE MATRIX OF THE CROSS-CORRELATION FUNCTION ESTIMATE

Consider the error of sample cross-correlation function (6) as

$$\hat{R}_{yu}(n)[N]-R_{yu}(n)=\hat{R}_{yu}(n)[N]-E\left\{\hat{R}_{yu}(n)[N]\right\}. \quad (40)$$

By substituting the output $y_{(k+n)}$ according to (2) we get (41), as shown at the bottom of the next page. Since $u_{(k)}$ and $d_{(k)}$ are deterministic, (41) gets

$$\hat{R}_{yu}(n)-R_{yu}(n)=\frac{1}{s(n)[N]}\sum_{k=1}^{N-n}\lambda^{(N-n-k)}\epsilon_{(k+n)}u_{(k)}. \quad (42)$$

The covariance matrix $\mathcal{Q}\in\mathbb{R}^{2(P+1)\times 2(P+1)}$ of the noise vector ζ (21) is essential for the generalized least squares method estimate (36) and hence to obtain the parameter estimate with minimal variance. This covariance matrix is formally defined as

$$\mathcal{Q}=E\left\{\zeta\zeta^{\mathrm{T}}\right\}, \quad (43)$$

which can be divided into submatrices as

$$\mathcal{Q}=\begin{pmatrix}\mathcal{Q}^{yuyu} & \mathcal{Q}^{yuyd}\\ \mathcal{Q}^{ydyu} & \mathcal{Q}^{ydyd}\end{pmatrix}=E\left\{\begin{pmatrix}\zeta^{yu}\zeta^{yu\mathrm{T}} & \zeta^{yu}\zeta^{yd\mathrm{T}}\\ \zeta^{yd}\zeta^{yu\mathrm{T}} & \zeta^{yd}\zeta^{yd\mathrm{T}}\end{pmatrix}\right\}. \quad (44)$$

The i-th row element and the j-th column element of submatrix $\mathcal{Q}^{yuyd}$ can be derived according to (42) and the definitions (19), (20) of vectors $\zeta^{\epsilon u}$, $\zeta^{\epsilon d}$ as

$$\begin{aligned}\mathcal{Q}_{ij}^{yuyd}[N]&=E\left\{\left(\hat{R}_{yu}(i)[N]-R_{yu}(i)[N]\right)\right.\\&\quad\left.\times\left(\hat{R}_{yd}(j)[N]-R_{yd}(j)[N]\right)\right\}\\&=\frac{1}{s(i)[N]s(j)[N]}E\left\{\sum_{k=1}^{N-i}\lambda^{(N-i-k)}\epsilon_{(k+i)}u_{(k)}\right.\\&\quad\left.\times\sum_{l=1}^{N-j}\lambda^{(N-j-l)}\epsilon_{(l+j)}d_{(l)}\right\}. \quad (45)\end{aligned}$$

By customizing a formula that can be found in [39], the identity [10]

$$E\left\{\sum_{k=1}^{m}X_k\sum_{l=1}^{n}Y_l\right\}=\sum_{k=1}^{m}\sum_{l=1}^{n}E\left\{X_kY_l\right\} \quad (46)$$

can be derived considering two general random vectors $X\in\mathbb{R}^{m\times 1}$ and $Y\in\mathbb{R}^{n\times 1}$. Applying the above identity,

equation (45) reduces to (47), as shown at the bottom of the next page.

Since ϵ is considered uncorrelated noise, its autocorrelation function is the Dirac delta function

$$R_{\epsilon\epsilon}(w) = \begin{cases} \sigma_\epsilon^2 & w = 0 \\ 0 & w \neq 0 \end{cases} \tag{48}$$

with the argument $w = k+i-l-j$ and the measure equal to the noise variance σ_ϵ^2 [32], [33].

Due to this property, the double summation in (47) $\forall i \geq j$ reduces to a single summation

$$\mathcal{Q}_{ij}^{yuyd}[N] = \frac{1}{s(i)[N]s(j)[N]}\sigma_\epsilon^2 \sum_{k=1}^{N-i} \lambda^{2(N-i-k)} u_{(k)} d_{(k+i-j)} \,. \tag{49}$$

In the case $i < j$ complementary to (49), the corresponding formula can be obtained as

$$\mathcal{Q}_{ij}^{yuyd}[N] = \frac{1}{s(i)[N]s(j)[N]}\sigma_\epsilon^2 \sum_{l=1}^{N-j} \lambda^{2(N-j-l)} d_{(l)} u_{(l+j-i)} \,. \tag{50}$$

One may notice that $\mathcal{Q}_{ji}^{yuyd} = \mathcal{Q}_{ij}^{ydyu}$, which implies that $\mathcal{Q}^{ydyu} = \left(\mathcal{Q}^{yuyd}\right)^{\mathrm{T}}$. The above steps can be taken to derive the remaining submatrices $\mathcal{Q}^{yuyu}$, $\mathcal{Q}^{ydyd}$ of (44).

Since the covariance matrix (44) is symmetric, submatrices $\mathcal{Q}^{yuyu}$, $\mathcal{Q}^{ydyd}$ are also symmetric, implying that $\mathcal{Q}_{ji}^{yuyu} = \mathcal{Q}_{ij}^{yuyu}$ and $\mathcal{Q}_{ji}^{ydyd} = \mathcal{Q}_{ij}^{ydyd}$.

The above equations can be considered as a generalization of the formulas from [10] for the case of exponentially weighted estimates of correlation functions.

A. RECURSIVE FORMULA FOR THE INVERSE

Since we are dealing with online identification, recursive relations are necessary to be derived to effectively update the covariance matrix $\mathcal{Q}$ based on the new input data, but more importantly to update its inverse $\mathcal{Q}^{-1}$.

Suppose that the elements of covariance submatrix $\mathcal{Q}^{\epsilon\epsilon ud}$ defined $\forall i \geq j$ by (49) can be obtained from $N+1$ available samples as (51), shown at the bottom of the next page.

The recursive formula can be obtained from (51) by substituting $\mathcal{Q}_{ij}^{yuyd}[N]$ according to (49) as (52), shown at the bottom of the next page. The recursive formula for the case $i < j$ complementary to (52) can be derived based on (50) as (53), shown at the bottom of the next page, which turns out to be equivalent to (52). The remaining recursive equations for updating submatrices $\mathcal{Q}^{yuyu}$, $\mathcal{Q}^{ydyd}$ can be derived analogously as (54) and (55), shown at the bottom of the next page.

To derive a recursive formula for updating the inverse $\mathcal{Q}^{-1}$, the Sherman-Morrison formula [40] will be exploited.

Considering an invertible matrix $\mathbf{A} \in \mathbb{R}^{\mathbf{N}\times\mathbf{N}}$ and column vectors $\mathbf{p}, \mathbf{r} \in \mathbb{R}^{\mathbf{N}\times 1}$ then formula

$$\left(\mathbf{A} + \mathbf{p}\mathbf{r}^{\mathrm{T}}\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{p}\mathbf{r}^{\mathrm{T}}\mathbf{A}^{-1}}{1 + \mathbf{r}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{p}} \tag{56}$$

holds for the rank-one update of the matrix inverse.

However, we have to slightly customize the Sherman-Morrison formula (56) to make it compatible with the problem of finding inverse $\mathcal{Q}^{-1}$ of the covariance matrix. Considering arbitrary column vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{\mathbf{N}\times 1}$ and matrices $\mathbf{D}, \mathbf{U}, \mathbf{V}, \mathbf{W} \in \mathbb{R}^{\mathbf{N}\times\mathbf{N}}$, we have the inversion lemma

$$\begin{aligned} &\left(\mathbf{V}^{-1}\mathbf{D}\mathbf{A}\mathbf{U}\mathbf{W}^{-1} + \mathbf{V}^{-1}\mathbf{u}\mathbf{v}^{\mathrm{T}}\mathbf{W}^{-1}\right)^{-1} \\ &\quad = \mathbf{W}\mathbf{U}^{-1}\left(\mathbf{A} + \mathbf{D}^{-1}\mathbf{u}\mathbf{v}^{\mathrm{T}}\mathbf{U}^{-1}\right)^{-1}\mathbf{D}^{-1}\mathbf{V} \,. \end{aligned} \tag{57}$$

Let $\mathbf{p} = \mathbf{D}^{-1}\mathbf{u}$, $\mathbf{r} = \left(\mathbf{U}^{-1}\right)^{\mathrm{T}}\mathbf{v}$ and $\mathbf{r}^{\mathrm{T}} = \mathbf{v}^{\mathrm{T}}\mathbf{U}^{-1}$, then according to (57) the Sherman-Morrison formula (56) can be customized as

$$\begin{aligned} &\left(\mathbf{V}^{-1}\mathbf{D}\mathbf{A}\mathbf{U}\mathbf{W}^{-1} + \mathbf{V}^{-1}\mathbf{u}\mathbf{v}^{\mathrm{T}}\mathbf{W}^{-1}\right)^{-1} \\ &\quad = \mathbf{W}\mathbf{U}^{-1}\mathbf{A}^{-1}\mathbf{D}^{-1}\mathbf{V} \\ &\qquad - \mathbf{W}\mathbf{U}^{-1}\frac{\mathbf{A}^{-1}\mathbf{D}^{-1}\mathbf{u}\mathbf{v}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{A}^{-1}}{1 + \mathbf{v}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{A}^{-1}\mathbf{D}^{-1}\mathbf{u}}\mathbf{D}^{-1}\mathbf{V} \,. \end{aligned} \tag{58}$$

For the inversion of covariance matrix $\mathcal{Q}$, matrix $\mathbf{A}$ represents the old covariance matrix $\mathcal{Q}[N]$ implying the dimension $\mathbf{N} = 2(P+1)$, and vectors $\mathbf{u} \in \mathbb{R}^{2(P+1)\times 1}$, $\mathbf{v} \in \mathbb{R}^{2(P+1)\times 1}$ have to correspond to update formulas (52), (53), (54), (55) such that

$$\mathbf{v} = \mathbf{u} = \begin{pmatrix} u_p \\ d_p \end{pmatrix} \sigma_\epsilon \,, \tag{59}$$

where vectors $u_p \in \mathbb{R}^{P+1\times 1}$ and $d_p \in \mathbb{R}^{P+1\times 1}$ are defined as

$$u_p = \begin{bmatrix} u_{(N+1)} & u_{(N)} & \dots & u_{(N+1-i)} & u_{(N+1-P)} \end{bmatrix}^{\mathrm{T}} \,, \tag{60a}$$

$$d_p = \begin{bmatrix} d_{(N+1)} & d_{(N)} & \dots & d_{(N+1-i)} & d_{(N+1-P)} \end{bmatrix}^{\mathrm{T}} \,. \tag{60b}$$

Scaling of the old covariance matrix $\mathcal{Q}[N]$ in update formulas (52), (53), (54), (55) is dependent on N and the indexes i, j. Therefore, matrices $\mathbf{D}, \mathbf{U}, \mathbf{V}, \mathbf{W}$ in (58) must have the diagonal structure

$$\mathbf{D} = \mathbf{U} = \lambda \,\mathrm{diag}\,(\mathbf{s}[N]) \,, \tag{61}$$

$$\mathbf{V} = \mathbf{W} = \mathrm{diag}\,(\mathbf{s}[N+1]) = \lambda\mathbf{D} + I \,, \tag{62}$$

$$\begin{aligned} \hat{R}_{yu}(n)[N] - R_{yu}(n) = \frac{1}{s(n)[N]} \Bigg[&\sum_{k=1}^{N-n} \lambda^{(N-n-k)} \left(\sum_{i=0}^{M_u} g_i^u[k+n] u_{(k+n-i)} + \sum_{i=0}^{M_d} g_i^d[k+n] d_{(k+n-i)} + \epsilon_{(k+n)} \right) u_{(k)} \\ &- E\left\{ \sum_{k=1}^{N-n} \lambda^{(N-n-k)} \left(\sum_{i=0}^{M_u} g_i^u[k+n] u_{(k+n-i)} + \sum_{i=0}^{M_d} g_i^d[k+n] d_{(k+n-i)} + \epsilon_{(k+n)} \right) u_{(k)} \right\} \Bigg] \,. \end{aligned} \tag{41}$$

where vectors $\mathbf{s}[N]$, $\mathbf{s}[N+1]$ get

$$\mathbf{s}[N] = \left[s(0)[N]\ s(1)[N]\ \ldots\ s(P)[N]\right]^{\mathrm{T}}\ , \tag{63a}$$

$$\mathbf{s}[N+1] = \left[s(0)[N+1]\ s(1)[N+1]\ \ldots\ s(P)[N+1]\right]^{\mathrm{T}}\ . \tag{63b}$$

Since $\mathbf{D}$ is diagonal, its inverse is easy to calculate.

The final formula for updating the inverse $\mathcal{Q}^{-1}$ can be derived according to (58) as

$$\mathcal{Q}^{-1}[N+1] = \mathbf{D}^{-1}\mathcal{Q}^{-1}[N]\mathbf{D}^{-1} - \mathbf{D}^{-1}\frac{\mathcal{Q}^{-1}[N]\mathbf{D}^{-1}\mathbf{v}\mathbf{v}^{\mathrm{T}}\mathbf{D}^{-1}\mathcal{Q}^{-1}[N]}{1+\mathbf{v}^{\mathrm{T}}\mathbf{D}^{-1}\mathcal{Q}^{-1}[N]\mathbf{D}^{-1}\mathbf{v}}\mathbf{D}^{-1}\ . \tag{64}$$

V. GRADIENT-BASED SOLUTION OF THE ESTIMATION PROBLEM

An emerging practical problem of the proposed online identification method is that at each iteration, new samples of signals $y_{(k)}$, $u_{(k)}$, $d_{(k)}$ are available, thus the regression system (13) changes completely in terms of full-rank update due to updated sample cross-correlation functions $\hat{R}_{yu}(n)$, $\hat{R}_{yd}(n)$, $\hat{R}_{ud}(n)$, $\hat{R}_{du}(n)$ and autocorrelation functions $\hat{R}_{uu}(n)$, $\hat{R}_{dd}(n)$. This fact hampers solving the estimation problem by the means of traditional online algorithms such as the recursive least squares method [11] suitable for rank-one update problems. Although we derived the recursive equations to effectively update the regression system itself and also update the covariance matrix $\mathcal{Q}$ and its inverse, there is still a need to re-evaluate the parameter estimate according to the closed form analytical solution (36) at each sample.

Due to the relatively large dimension $2(P+1)\times(M_u+M_d+2)$ of the estimation problem, the complexity of calculating the inverse of $M_u+M_d+2\times M_u+M_d+2$ Hessian matrix in (36) is significant, hence it may be even infeasible to perform it in real time. Therefore, we propose to rather use a suitable numeric optimization method to find the minimum of the quadratic form (33), in particular, the conjugate gradient method [41] will be adopted.

Consider an unconstrained minimization problem of a multivariate scalar cost function $f(\mathbf{x}) : \mathbb{R}^{n\times 1} \to \mathbb{R}$ as $\min f(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^{n\times 1}$ and $n = M_u + M_d + 2$. To disambiguate the introduced notation, the gradient of the general cost function $f(\mathbf{x})$ will be denoted as $\mathbf{g}(\mathbf{x}) : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times 1}$ and for the Hessian, notation $\mathbf{H}(\mathbf{x}) : \mathbb{R}^{n\times 1} \to \mathbb{R}^{n\times n}$ will be used.

The algorithm of conjugate gradient method is stated as follows [43]. Considering i is the iteration number, the formula to update the approximation of the optimal solution gets the recursive form

$$\mathbf{x}_i = \mathbf{x}_{i-1} + \mathbf{s}_i\delta_i\ , \tag{65}$$

where $\mathbf{x}_{i-1}$ is the last and $\mathbf{x}_i$ is the new approximation of the optimal solution, $\mathbf{s}_i \in \mathbb{R}^{n\times 1}$ is the direction vector and $\delta_i \in \mathbb{R}$ is the step size.

The optimal step size δ_i can be determined as

$$\delta_i = -\frac{\mathbf{s}_i^{\mathrm{T}}\mathbf{g}(\mathbf{x}_{i-1})}{\mathbf{s}_i^{\mathrm{T}}\mathbf{H}(\mathbf{x}_{i-1})\mathbf{s}_i}\ . \tag{66}$$

The direction vector $\mathbf{s}_{i+1}$ follows the recursive formula

$$\mathbf{s}_{i+1} = \mathbf{s}_i\gamma_i - \mathbf{g}(\mathbf{x}_i)\ , \tag{67}$$

where the scalar coefficient $\gamma_i \in \mathbb{R}$ can be determined as

$$\gamma_i = \frac{\mathbf{g}(\mathbf{x}_i)^{\mathrm{T}}\mathbf{g}(\mathbf{x}_i)}{\mathbf{g}(\mathbf{x}_{i-1})^{\mathrm{T}}\mathbf{g}(\mathbf{x}_{i-1})}\ . \tag{68}$$

This iterative algorithm starts with $\mathbf{s}_0 = -\mathbf{g}(\mathbf{x}_0)$.

$$\mathcal{Q}_{ij}^{yuyd}[N] = \frac{1}{s(i)[N]s(j)[N]}\sum_{k=1}^{N-i}\sum_{l=1}^{N-j}\lambda^{(N-i-k)}\lambda^{(N-j-l)}R_{\epsilon\epsilon}(k+i-l-j)u_{(k)}d_{(l)} \tag{47}$$

$$\mathcal{Q}_{ij}^{yuyd}[N+1] = \frac{1}{s(i)[N+1]s(j)[N+1]}\left(\lambda^2\sigma_\epsilon^2\sum_{k=1}^{N-i}\left[\lambda^{2(N-i-k)}u_{(k)}d_{(k+i-j)}\right] + \sigma_\epsilon^2 u_{(N+1-i)}d_{(N+1-j)}\right) \tag{51}$$

$$\mathcal{Q}_{ij}^{yuyd}[N+1] = \frac{1}{s(i)[N+1]s(j)[N+1]}\left(\lambda^2 s(i)[N]s(j)[N]\mathcal{Q}_{ij}^{yuyd}[N] + \sigma_\epsilon^2 u_{(N+1-i)}d_{(N+1-j)}\right) \tag{52}$$

$$\mathcal{Q}_{ij}^{yuyd}[N+1] = \frac{1}{s(i)[N+1]s(j)[N+1]}\left(\lambda^2 s(i)[N]s(j)[N]\mathcal{Q}_{ij}^{yuyd}[N] + \sigma_\epsilon^2 d_{(N+1-j)}u_{(N+1-i)}\right) \tag{53}$$

$$\mathcal{Q}_{ij}^{yuyu}[N+1] = \frac{1}{s(i)[N+1]s(j)[N+1]}\left(\lambda^2 s(i)[N]s(j)[N]\mathcal{Q}_{ij}^{yuyu}[N] + \sigma_\epsilon^2 u_{(N+1-i)}u_{(N+1-j)}\right) \tag{54}$$

$$\mathcal{Q}_{ij}^{ydyd}[N+1] = \frac{1}{s(i)[N+1]s(j)[N+1]}\left(\lambda^2 s(i)[N]s(j)[N]\mathcal{Q}_{ij}^{ydyd}[N] + \sigma_\epsilon^2 d_{(N+1-i)}d_{(N+1-j)}\right) \tag{55}$$

In the case of minimizing the quadratic form (33) of the parameter estimation problem, the gradient corresponds to (34) and the Hessian can be determined according to (35). The rationale for the conjugate gradient method is that the algorithm can be terminated right after n iterations obtaining the exact solution equivalent to (36). The termination can be triggered if the magnitude of the gradient decreases below a certain threshold, i.e. if $\mathbf{g}(\mathbf{x}_i)^{\mathrm{T}}\,\mathbf{g}(\mathbf{x}_i) < \kappa$.

As was already mentioned, for each new sample, the optimal solution of the parameter estimation problem is expected to change (drift), hence the conjugate gradient algorithm has to be iterated again. However, the initial guess $\mathbf{x}_0[N{+}1]$ can be chosen equal to the optimal solution obtained from the last sample $\mathbf{x}_n[N]$ to speed up the convergence.

VI. APPROXIMATE SOLUTION TO THE MATRIX INVERSE

In this section, we will propose an alternative approach to find a solution to the parameter estimation problem based on the closed form (36). The main motivation is that instead of directly computing the inverse $\mathbf{H}^{-1}$ of the Hessian matrix (35) by the means of standard algorithms of linear algebra at each sample, the matrix inverse will be approximated using a dedicated iterative method, as discussed below.

Consider a matrix function $F(\mathbf{X}) : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ representing the residual of the approximation of the inverse $\mathbf{H}^{-1}$ by matrix $\mathbf{X} \in \mathbb{R}^{n\times n}$ as

$$F(\mathbf{X}) = \mathbf{X}^{-1} - \mathbf{H}\,, \tag{69}$$

where $n = M_u + M_d + 2$.

To find the root of the residual function (69) i.e. the solution to $F(\mathbf{X}^\star) = \mathbf{0}$ that satisfies $\mathbf{X}^\star = \mathbf{H}^{-1}$, the Newton-Raphson method generalized for matrix functions can be conveniently adopted. The outlined combination of using the Newton-Raphson method to approximate the matrix inverse is referred to as the Schulz method in the literature [44].

The Schulz method iteration [46] gets

$$\mathbf{X}_{i+1} = \mathbf{X}_i\,(2I - \mathbf{H}\mathbf{X}_i)\ . \tag{70}$$

In [46], it was proved that the magnitudes of all the eigenvalues of $\mathbf{H}\mathbf{X}_0 - I$ must be less than 1 to ensure the convergence of the algorithm (70), what can also be noted as

$$\rho(\mathbf{H}\mathbf{X}_0 - I) < 1\,, \tag{71}$$

where operator $\rho(\cdot)$ represents the magnitude of the largest eigenvalue, the so-called spectral radius.

A. STANDARD CHOICE OF THE INITIAL GUESS

The initial guess $\mathbf{X}_0$ in the Schulz method is typically chosen as [44]

$$\mathbf{X}_0 = a\mathbf{H}^{\mathrm{T}}\,, \tag{72}$$

where $a \in \mathbb{R}$.

Substituting (72) into the convergence criterion (71) yields

$$0 < \rho\left(a\mathbf{H}\mathbf{H}^{\mathrm{T}} - I\right) < 1\,. \tag{73}$$

At this point, it is crucial that the inequality [47], [48]

$$\rho(\mathbf{A} + \mathbf{B}) \le \rho(\mathbf{A}) + \rho(\mathbf{B}) \tag{74}$$

holds for the spectral radius of the sum of two permutable matrices $\mathbf{A}$, $\mathbf{B}$, i.e. the matrices satisfying $\mathbf{AB} = \mathbf{BA}$. Since the spectral radius of the identity matrix is $\rho(I) = 1$, one can reshape criterion (73) according to (74) as

$$0 < a < \frac{2}{\rho\left(\mathbf{H}\mathbf{H}^{\mathrm{T}}\right)}\,. \tag{75}$$

However, to determine a that satisfies (75), it is not necessary to calculate the exact spectral radius $\rho\left(\mathbf{H}\mathbf{H}^{\mathrm{T}}\right)$ in practice, since the so-called Gershgorin circle theorem can be assumed instead.

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a square matrix and R_i the sum of the absolute values of the non-diagonal entries in the i-th row i.e. $R_i = \sum_{j\neq i} |\mathbf{A}_{ij}|$. Then every eigenvalue of $\mathbf{A}$ lies within at least one of the Gershgorin discs $D(\mathbf{A}_{ii}, R_i) \in \mathbb{C}$ with the center at $\mathbf{A}_{ii}$ and the radius R_i [49]. This theorem implies that, for the spectral radius of $\mathbf{A}$, the following inequality holds

$$\rho(\mathbf{A}) \le \max\left\{\sum_{j=1}^{n} |\mathbf{A}_{ij}|\ \forall i = 1\ldots n\right\}\,. \tag{76}$$

Considering (76) yields the condition (75) modified as [50]

$$0 < a < \frac{2}{\max\left\{\sum_{j=1}^{n} \left|\left(\mathbf{H}\mathbf{H}^{\mathrm{T}}\right)_{ij}\right|\ \forall i = 1\ldots n\right\}}\,. \tag{77}$$

B. RECURSIVE CHOICE OF THE INITIAL GUESS

A more effective strategy to choose the initial guess in the Schulz method is to exploit the inverse of the Hessian $\mathbf{H}^{-1}[N]$ obtained from the last sample as

$$\mathbf{X}_0 = a\mathbf{H}^{-1}[N]\,, \tag{78}$$

while anticipating that the inter-sample perturbation of the Hessian is not too large.

It has to be ensured that the initial guess (78) satisfies the convergence criterion (71) as

$$0 < \rho\left(a\mathbf{H}[N{+}1]\mathbf{H}^{-1}[N] - I\right) < 1\,. \tag{79}$$

As inequality (74) holds, the above criterion reduces to

$$0 < a < \frac{2}{\rho\left(\mathbf{H}[N{+}1]\mathbf{H}^{-1}[N]\right)}\,. \tag{80}$$

Since the Gershgorin circle theorem (76) holds, condition (80) gets modified as

$$0 < a < \frac{2}{\max\left\{\sum_{j=1}^{n} \left|\left(\mathbf{H}[N{+}1]\mathbf{H}^{-1}[N]\right)_{ij}\right|\ \forall i = 1\ldots n\right\}}\,. \tag{81}$$

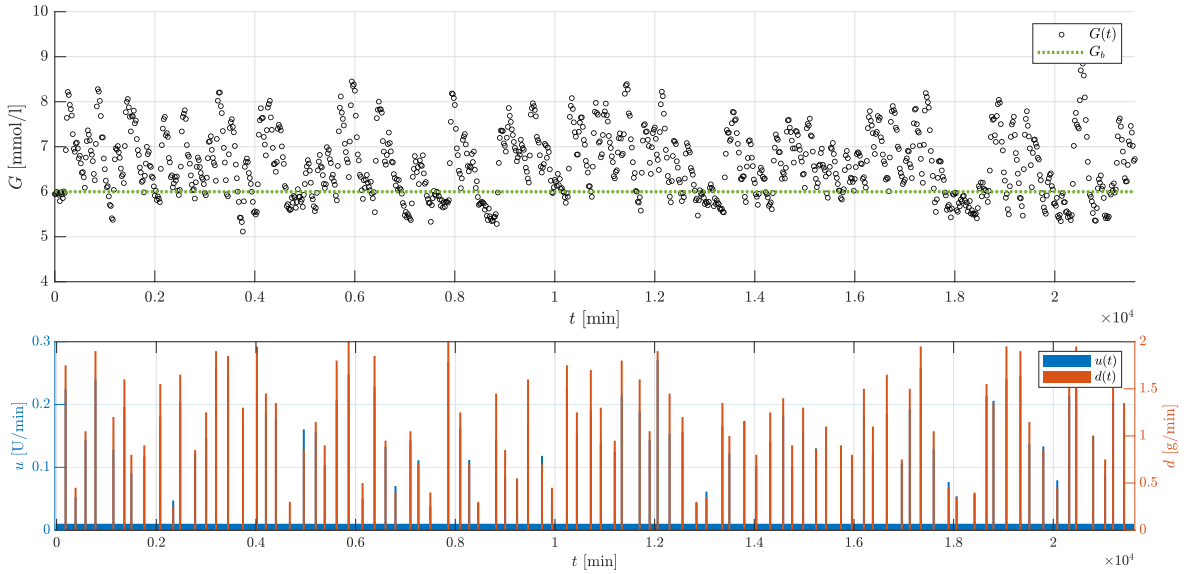


FIGURE 1. Input-output diabetic dataset obtained without the influence of parametric variability.

C. ALGORITHM TERMINATION

To trigger the termination of iterative algorithm (70) a stopping criterion is required. Multiplying residual function (69) by $\mathbf{X}$ yields matrix function $\mathbf{E}(\mathbf{X})$ as

$$\mathbf{E}(\mathbf{X}) = I - \mathbf{H}\mathbf{X}\,. \tag{82}$$

The convergence can be quantified by the scalar quadratic criterion q

$$q(\mathbf{X}_i) = \sum_{p=1}^{n}\sum_{r=1}^{n}\mathbf{E}(\mathbf{X}_i)_{pr}^2\,. \tag{83}$$

The stopping condition can be formulated as $q(\mathbf{X}_i) < \left(10^{-20}, 10^{-10}\right)$ depending on the desired approximation accuracy.

VII. SIMULATION EXPERIMENT

As outlined in the title, the target application domain of the proposed online identification algorithm is the parameter-varying empirical model of glycemia dynamics in patients with type 1 diabetes. To validate the design and effectiveness of the proposed robust online identification, in this section, the virtual patient diabetic data generated by a complex physiology-based simulation model will be considered.

The glycemia response $G(t)$ [mmol/l] for this experiment was obtained by the in-silico approach, using the nonlinear simulation model discussed in [51] and [52]. The mean population parameters available in [51] were adopted, while some adjustments had to be made in order to simulate the metabolic specifics of type 1 diabetes [53]. 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.

Data acquisition experiments were designed to mimic the standard insulin treatment of a type 1 diabetic subject during the 15-day period with a total number of 90 meals. Virtual continuous glucose monitoring data were sampled with $T_s = 20$ min, while the total duration of the experiment can be determined as $15 \times 60 \times 24$ min, which implies the number of samples $N = 1080$. The measurements were distorted by the additive uncorrelated zero-mean noise with the standard deviation $\sigma_\epsilon = 0.1$ mmol/l. A typical diabetic dataset includes information on administered insulin boluses and carbohydrate intake in the form of a diabetic diary. During conventional insulin therapy, the insulin dose is determined according to the so-called bolus calculator rule (see [54] and the references therein).

The first dataset obtained under the assumption that the simulation model parameters are fixed to their mean population values without the influence of parametric variability is depicted in Figure 1.

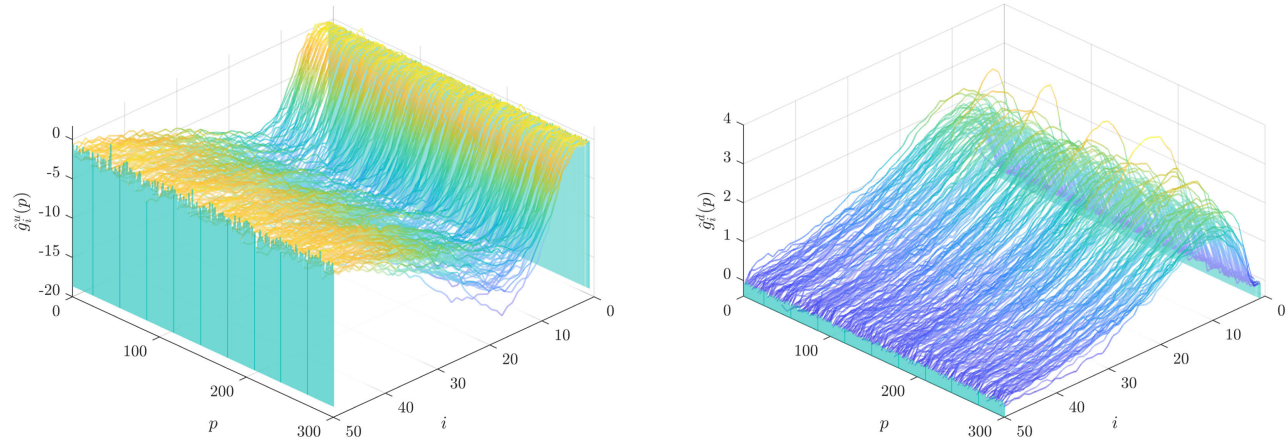
The prediction performance of the identified models will be quantified by metric

$$Q = \frac{1}{N-P}\sum_{k=P}^{N}\left[\hat{y}_{(k)} - y_{(k)}\right]^2\,, \tag{84}$$

where $\hat{y}_{(k)}$ is the predicted model output and $y_{(k)}$ is the measured output.

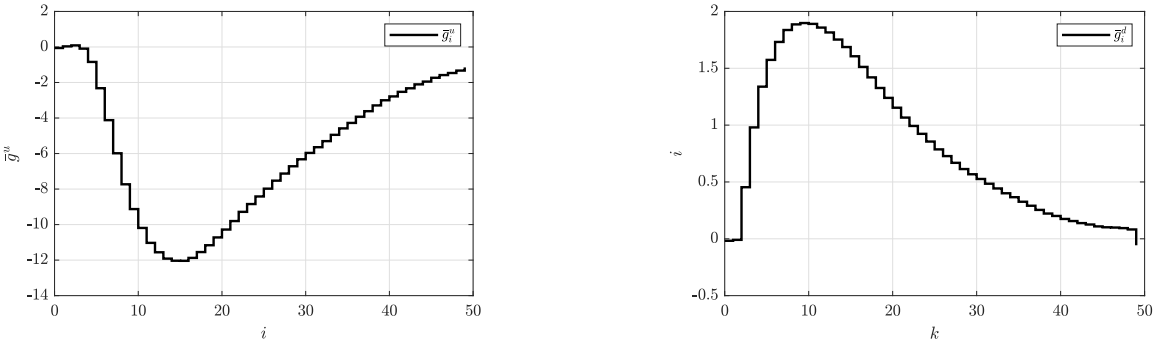
A. MODELING THE INTER-SUBJECT VARIABILITY

To model the inter-subject variability and to estimate the essential covariance matrix $\hat{\Psi}$ and the mean-population parameter vector $\bar{g}$ for the sake of regularization and robustification of the estimate in the terms of section III-C, a special experiment was designed. The population of virtual diabetic subjects had to be randomly generated, assuming the normal distribution of all parameters of the simulation model with the available mean-population value $\bar{p}_i$ and standard deviation σ_{p_i} determined according to the uniform coefficient of variation strategy. The coefficient of variation c_v will be



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

FIGURE 2. Non-regularized estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u(p)$ and the carbohydrate intake effect submodel $\hat{g}^d(p)$ obtained from $n_p = 300$ virtual diabetic subjects.



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

FIGURE 3. Estimated mean-population impulse responses of the insulin administration effect submodel $\bar{g}^u$ and the carbohydrate intake effect submodel $\bar{g}^d$.

defined as a fixed ratio $c_v = \frac{\sigma_{p_i}}{\bar{p}_i} = 0.15$ for all model parameters.

For each of $n_p = 300$ virtual diabetic subjects, the glycemia response was generated and the nonregularized estimate of impulse responses assuming $\alpha = 0$, $\beta = 0$ and no forgetting $\lambda = 1$ was carried out, while the results for the final sample $N = 1080$ can be seen in Figure 2.

This dataset was consequently used to estimate the mean-population impulse response $\bar{g}$ according to (26) and to determine the covariance matrix $\hat{\Psi}$ according to (27). The obtained $\bar{g}$ is plotted in Figure 3, whereas the estimated $\hat{\Psi}$ is visualized in Figure 4 as a colormap. In Figure 4a, one can notice that the variance of the insulin administration effect impulse response is much higher during the peak than during the rise or decay phase, whereas the submatrix corresponding to carbohydrate intake in Figure 4b shows higher variances at the beginning of the response.

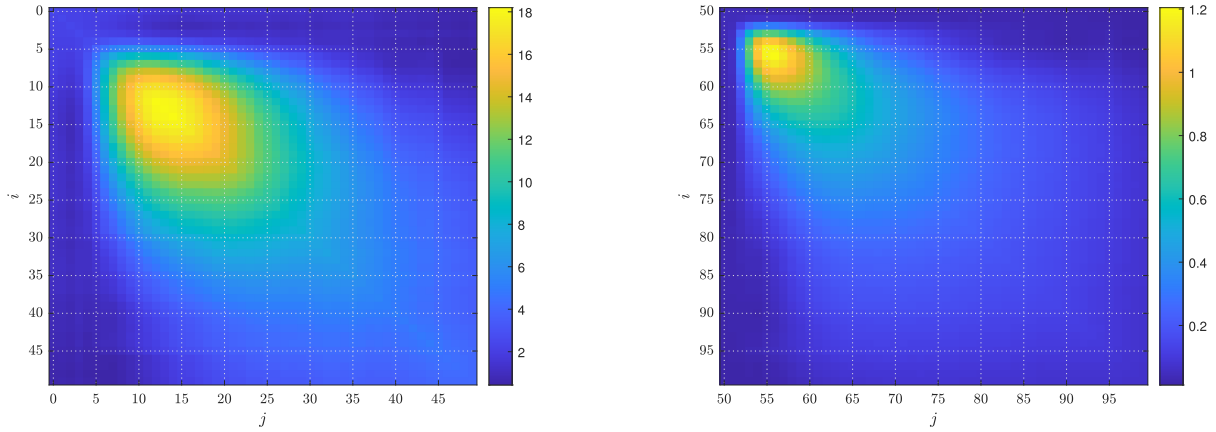
B. MODELING THE INTRA-SUBJECT VARIABILITY

To involve the mechanism of intra-subject time variability into the in-silico experiment, chosen parameters of the simulation model were considered and implemented as time-varying. Since there are actually no studies available that would report on the in-vivo experience or publish experimental data that describe the parametric variability of the concerned simulation model, the evolution curves of these parameters will be designed empirically. The list of chosen physiology-based parameters, which will be considered time-varying, is summarized in Table 1. For a more comprehensive description and physiological interpretation of the chosen parameters, see [51] and the references therein.

The designed curves representing the deviations of the corresponding parameters $p_i(t)$ from the nominal values $\bar{p}_i$ divided by $\bar{p}_i$ to obtain dimensionless quantities according to $\Delta p_i(t) = \frac{p_i(t) - \bar{p}_i}{\bar{p}_i}$ are summarized in Figure 5.

By simulating the model under the influence of empirically designed parametric variabilities from Figure 5, another dataset was obtained as documented in Figure 6. This dataset will be used to demonstrate and benchmark the adaptive capabilities of the proposed identification algorithm.

Based on the results obtained by studying $n_p = 300$ virtual diabetic subjects, Figure 7 shows the covariance matrix $\hat{\Phi}$ of the difference between two consecutive parameter estimates estimated according to (32) representing the inter-sample



(a) $\hat{\Psi}_{ij}\ \forall i < M_u, j < M_u$ corresponding to $\hat{\Psi}^{uu}$

(b) $\hat{\Psi}_{ij}\ \forall i \geq M_u, j \geq M_u$ corresponding to $\hat{\Psi}^{dd}$

FIGURE 4. Entries of the estimated covariance matrix $\hat{\Psi}$.

TABLE 1. Chosen time-varying parameters of the simulation model.

symbol	unit	parameter name or description of effect	subsystem
k_1	min^{-1}	rate parameter	glucose subsystem
k_2	min^{-1}	rate parameter	glucose subsystem
m_1	min^{-1}	rate parameter	insulin kinetics subsystem
m_2	min^{-1}	rate parameter	insulin kinetics subsystem
k_{gri}	min^{-1}	rate of grinding	glucose intestinal absorption subsystem
k_{abs}	min^{-1}	rate constant of gastric emptying	glucose intestinal absorption subsystem
k_{p3}	$\frac{\mathrm{mg/kg/min}}{\mathrm{pmol/l}}$	affects amplitude of insulin action on the liver	EGP subsystem
k_i	min^{-1}	affects delay between insulin signal and insulin action	EGP subsystem
p_{2U}	min^{-1}	rate of insulin action on the peripheral glucose utilization	glucose utilization subsystem

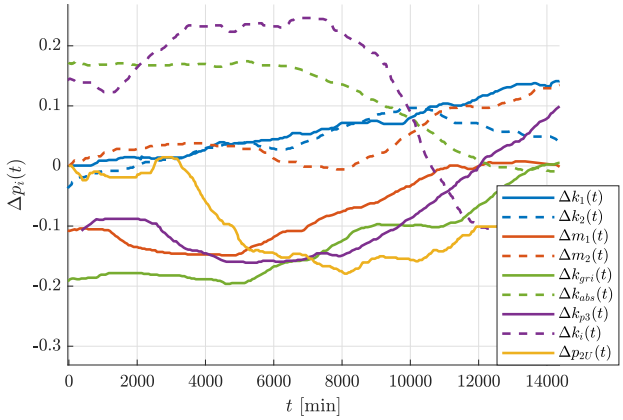


FIGURE 5. Scale-invariant relative representation of evolution of the chosen time-varying parameters of the simulation model.

parametric variability. This figure shows that the variability of insulin action is more prominent than the carbohydrate intake response, while it affects primarily the peak time and the amplitude.

The entries of the inverse of the covariance matrix $\mathcal{Q}^{-1}[N]$ corresponding to the final sample of the dataset from Figure 1, which was obtained recursively according to (64), are visualized as a colormap in Figure 8. One can notice a quasi-diagonal structure of $\mathcal{Q}^{-1}[N]$, different scales of its submatrices, and also the presence of slightly lower values of diagonal entries with an increasing index.

C. MAIN RESULTS

Finally, the estimated impulse responses under various configurations will be presented and discussed.

Concerning the tuning of the proposed identification algorithm, the number of lags was chosen as $P = 100$ and the lengths of the identified impulse responses were adjusted as $M_u = 49$, $M_d = 49$. The forgetting factor for the first non-adaptive scenario was set to $\lambda = 1$ and for the second scenario that assumed the effect of parametric variability and online adaptation of impulse responses, it was chosen as $\lambda = 0.995$.

The not regularized scenario with $\alpha = 0$, $\beta = 0$ and $\lambda = 1$ is visualized in Figure 9, while it was obtained by processing the parameter-invariant dataset from Figure 1. It was demonstrated that if no regularization is applied, the coefficients of impulse response are disorganized, hence less valid, and definitely not physiology compliant.

By applying the first regularization term with $\alpha = 5 \times 10^2$ and omitting the second regularization term as $\beta = 0$, the sequence of impulse responses documented in Figure 10 was

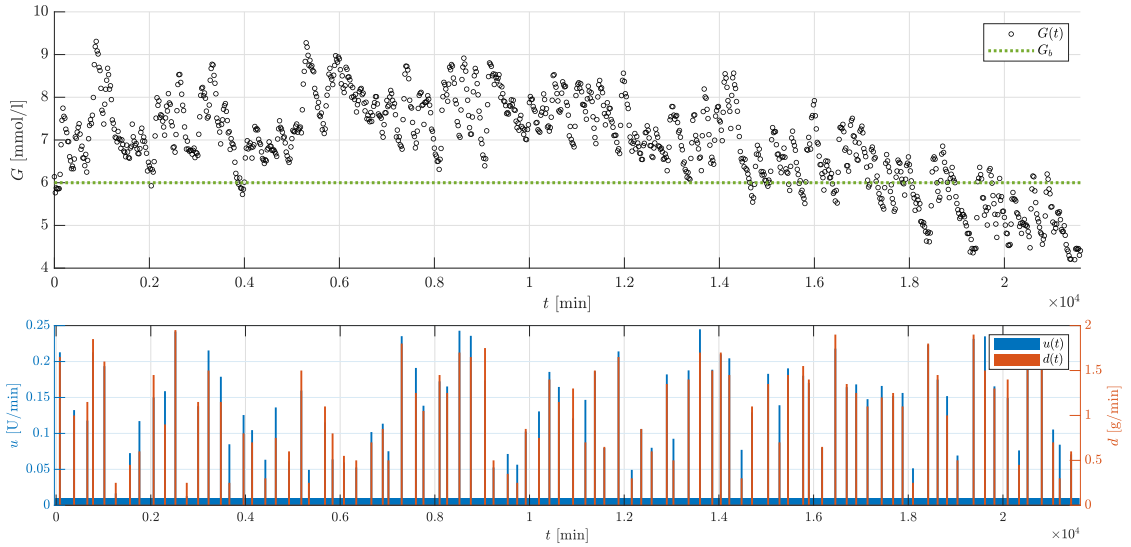
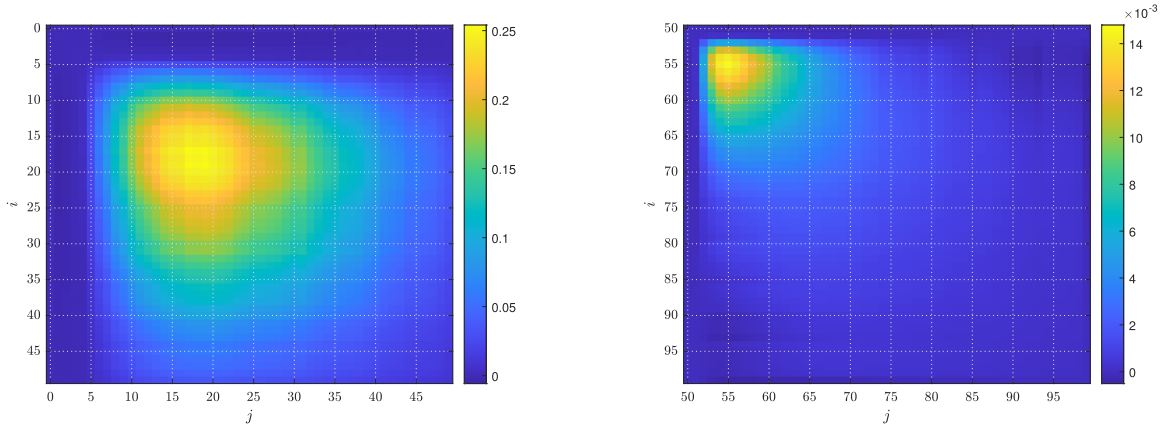


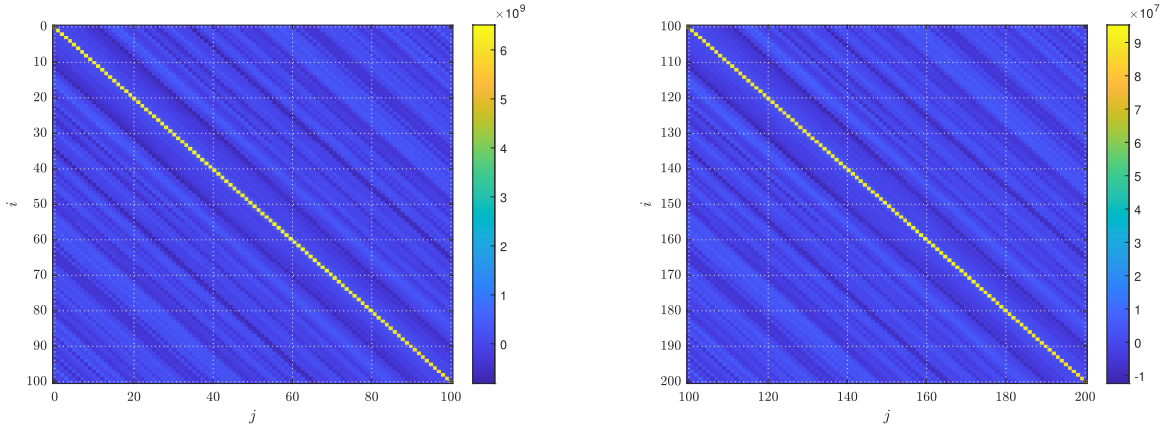
FIGURE 6. Input-output diabetic dataset obtained under the assumption of time-varying parameters.



(a) $\hat{\Phi}_{ij}\ \forall i < M_u, j < M_u$ corresponding to $\hat{\Phi}^{uu}$

(b) $\hat{\Phi}_{ij}\ \forall i \geq M_u, j \geq M_u$ corresponding to $\hat{\Phi}^{dd}$

FIGURE 7. Entries of the estimated covariance matrix $\hat{\Phi}$.



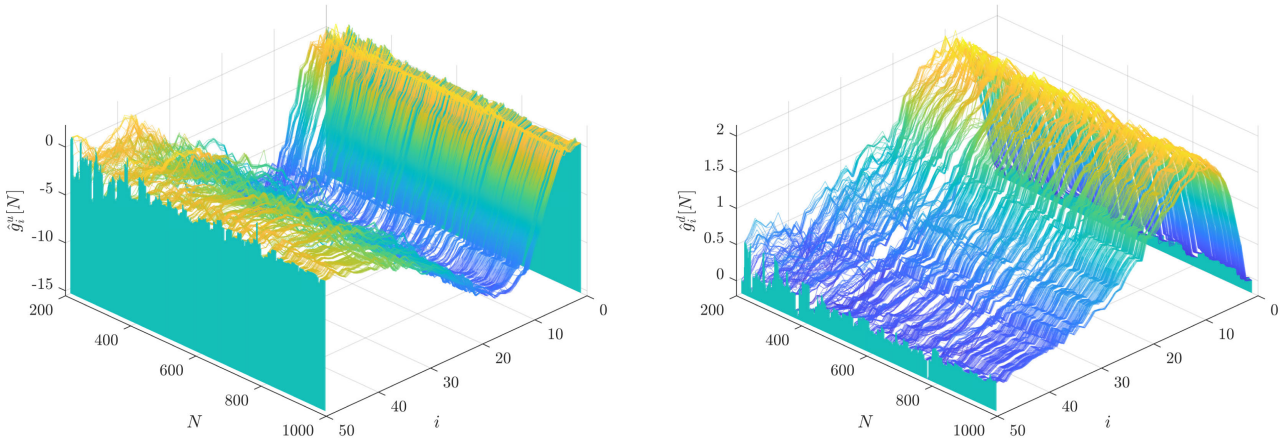
(a) $\left(\mathcal{Q}^{-1}\right)_{ij}\ \forall i < P, j < P$ corresponding to the insulin administration input

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

FIGURE 8. The elements of the inverse of the covariance matrix $\mathcal{Q}^{-1}[N]$.

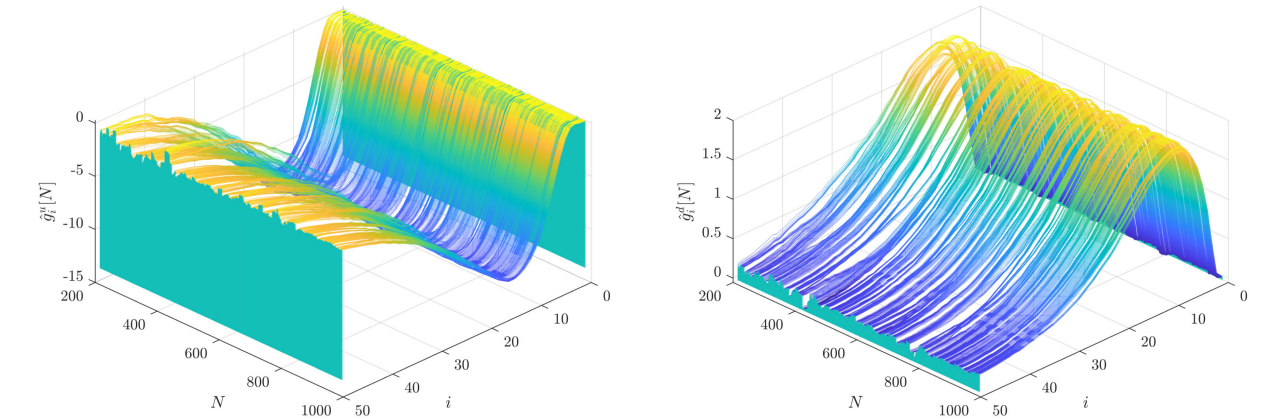
obtained by processing the parameter-invariant dataset from Figure 1.

Assuming both regularization strategies by setting $\alpha = 5 \times 10^2$, $\beta = 1 \times 10^{-3}$ in order to obtain a more robust and



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

FIGURE 9. Not regularized estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the parameter-invariant dataset.



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

FIGURE 10. Regularized $\alpha \neq 0$, $\beta = 0$ estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the parameter-invariant dataset.

physiology-compliant estimate, satisfying results could be obtained by processing the parametric-invariant dataset from Figure 1 as summarized in Figure 11. It can be observed that there was almost no drift of parameters, yet small fluctuations were caused by the nonlinearity of the simulation model and the related adaptation with respect to the change of operating point.

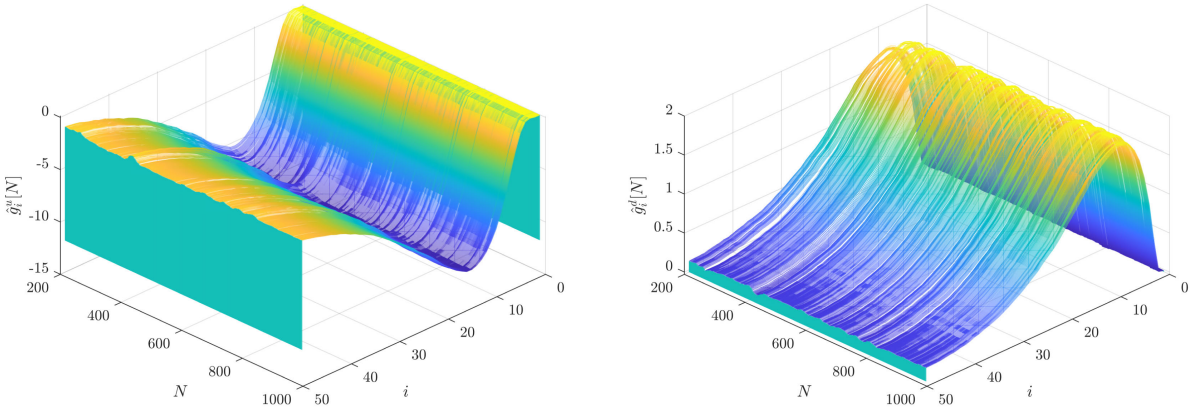
The presented results show that the impulse responses represent a stable system with aperiodic nature. The estimated coefficients keep their sign uniform, which implies that the estimate is physiology compliant. Additionally, a slower and longer lasting effect of insulin administration can be observed with a peak at approximately 300 minutes compared to a faster and shorter lasting effect of carbohydrate intake, which peaked roughly at 200 minutes.

The next experiment concerns the online identification performed to adapt the estimated impulse responses due to the influence of intra-subject parametric variability. Assuming both regularization strategies with the corresponding weights chosen as $\alpha = 5 \times 10^2$, $\beta = 1 \times 10^{-3}$ and the forgetting factor $\lambda = 0.995$, by processing the parameter-varying dataset from Figure 6 we obtained the results summarized in Figure 12. Significant drift and adaptation of estimated impulse responses can be observed, as the chosen physiology-based parameters of the simulation model (see Table 1) varied according to Figure 5.

The identification was repeated considering a smaller forgetting factor $\lambda = 0.95$, which implies that the forgetting effect will be stronger in this case. Figure 13 then documents that both impulse responses evolve very rapidly with a low consistency, hence their overall validity cannot be considered satisfactory.

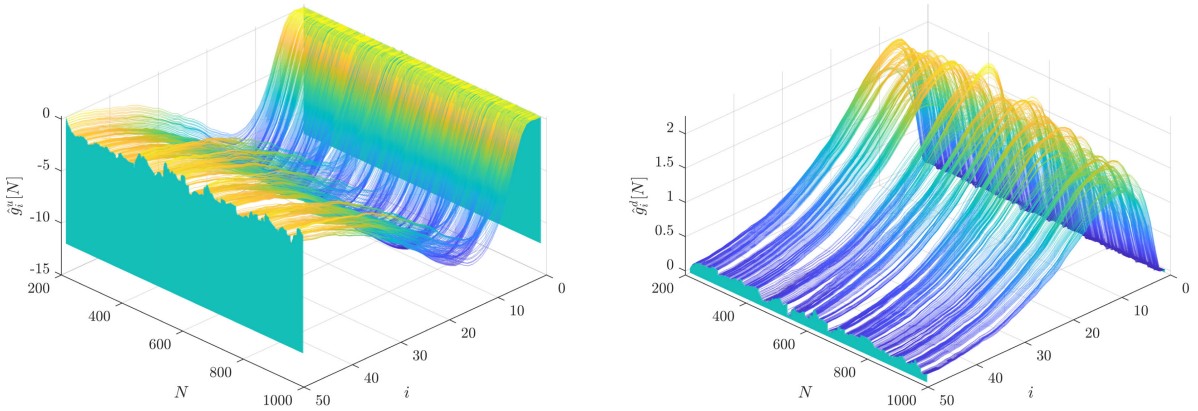
To validate the performance of the proposed identification algorithm with respect to the presence of outliers in the input data, the glycemia measurements from Figure 6 will be distorted by considering multiple randomly scattered outliers (invalid measurements) as can be seen in Figure 14.

Assuming both regularization strategies with the corresponding weights chosen as $\alpha = 5 \times 10^2$, $\beta = 1 \times 10^{-3}$ and the forgetting factor $\lambda = 0.995$, by processing the dataset from



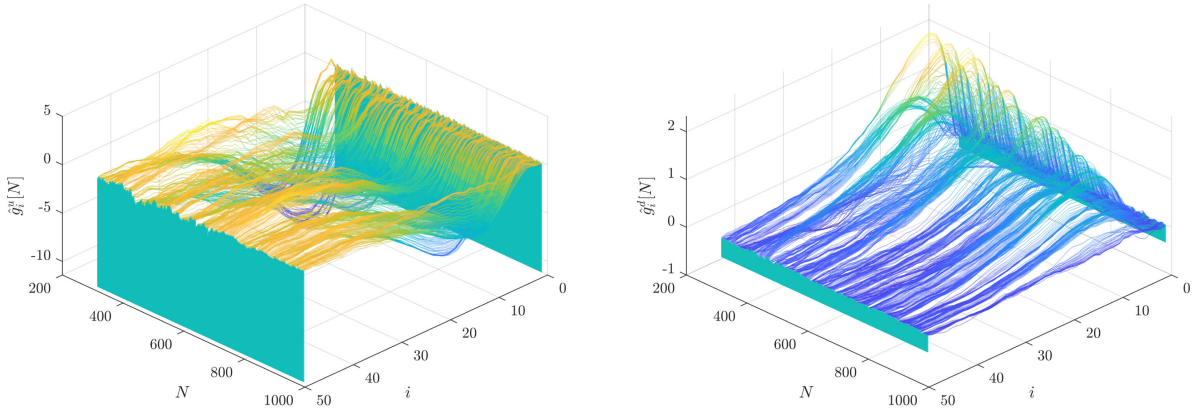
(a) Insulin administration effect submodel (b) Carbohydrate intake effect submodel

FIGURE 11. Fully regularized $\alpha \neq 0$, $\beta \neq 0$ estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the parameter-invariant dataset.



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

FIGURE 12. Online adaptive estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the parameter-varying dataset while considering $\lambda = 0.995$.



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

FIGURE 13. Online adaptive estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the parameter-varying dataset while considering $\lambda = 0.95$.

Figure 14, the impulse responses were identified as visualized in Figure 15.

It can be concluded that the outliers affected the estimated model, but due to the applied regularization, larger

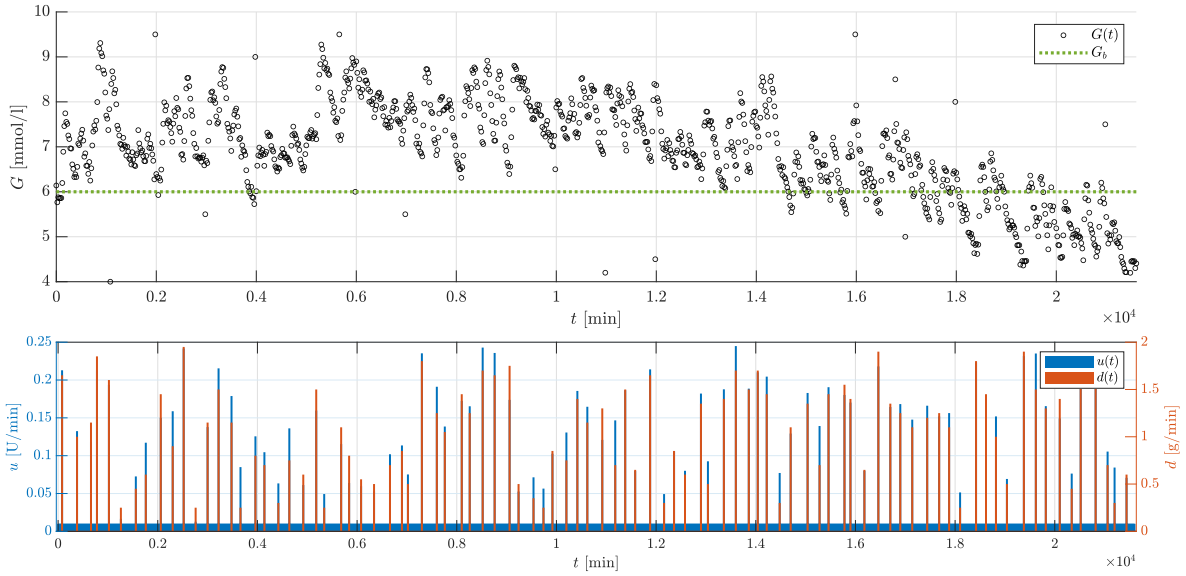
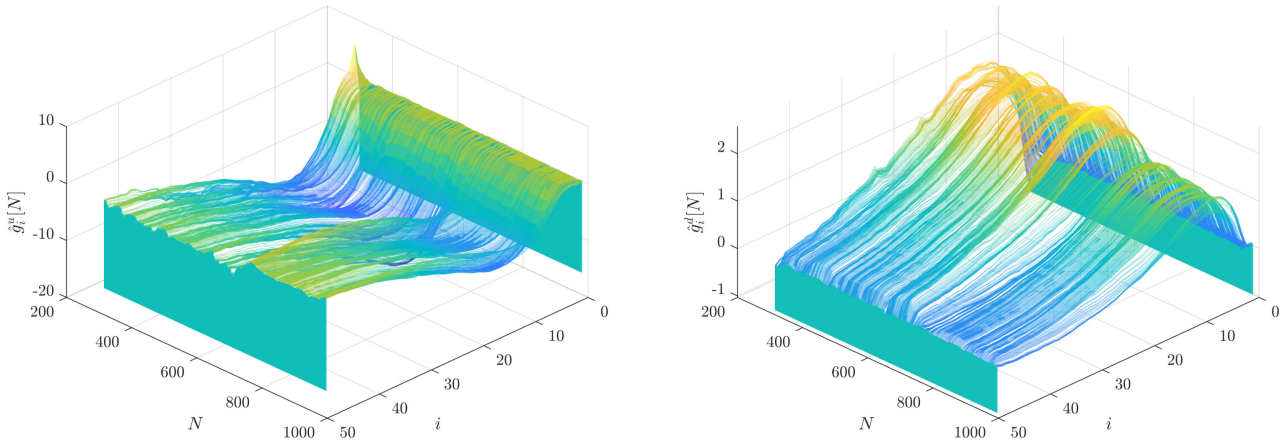


FIGURE 14. Input-output diabetic dataset distorted by the outliers.



(a) Insulin administration effect submodel

(b) Carbohydrate intake effect submodel

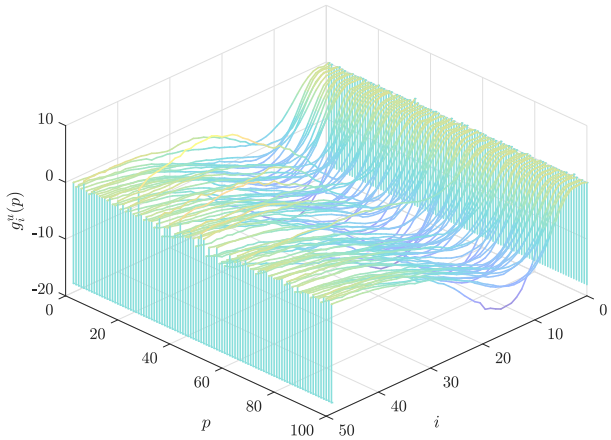
FIGURE 15. Online adaptive estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the dataset distorted by the outliers.

perturbations of the estimated impulse responses were suppressed, and the model remained relatively valid.

To assess how the proposed identification method performs across a broad demographic spectrum of type 1 diabetic patients in terms of intersubject variability, another experiment is designed and evaluated. To properly model the inter-subject variability, a new population of virtual diabetic subjects was randomly generated assuming the normal distribution of all parameters of the simulation model while considering known mean-population value $\bar{p}_i$ and standard deviation σ_{p_i} following the uniform coefficient of variation $c_v = \frac{\sigma_{p_i}}{\bar{p}_i} = 0.15$ for each of the model parameters.

For each of $n_p = 100$ virtual diabetic subjects, the glycemia response was generated and the estimation of impulse responses assuming both regularizations with weights $\alpha = 5 \times 10^2$, $\beta = 1 \times 10^{-3}$ and the forgetting factor $\lambda = 0.995$ was performed. Considering the final sample $N = 1080$, we obtained the impulse responses as visualized in Figure 16. These impulse responses show significant variations of the magnitudes and times of the peaks across the population, while the majority of responses can be considered compliant with physiology.

Besides the basic assessment of impulse responses, the estimated models were validated by predicting glycemia individually for each subject of the dataset, while the average value of performance criterion (84) resulted in $\bar{Q} = 0.1097$ and its variance was $\sigma_Q^2 = 0.0354$, which can be considered satisfactory. Unfortunately, further visualization of the evolution of impulse responses with respect to number of samples N and the corresponding prediction of glycemia for each subject of the set is unfeasible.



(a) Insulin administration effect submodel

(b) Carbohydrate intake effect submodel

FIGURE 16. Estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u(p)$ and the carbohydrate intake effect submodel $\hat{g}^d(p)$ obtained from $n_p = 100$ virtual subjects across broad demographic spectrum of type 1 diabetic patients generated in terms of intersubject variability.

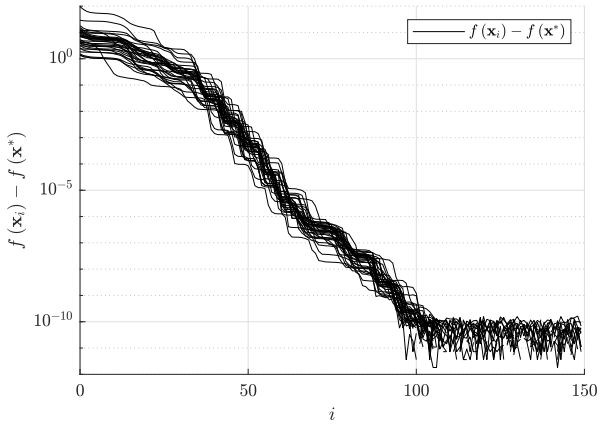


FIGURE 17. Convergence of cost function $f(\mathbf{x}_i)$ obtained through the iterations i of the conjugate gradient method taken relatively to the explicit analytical solution $\mathbf{x}^*$.

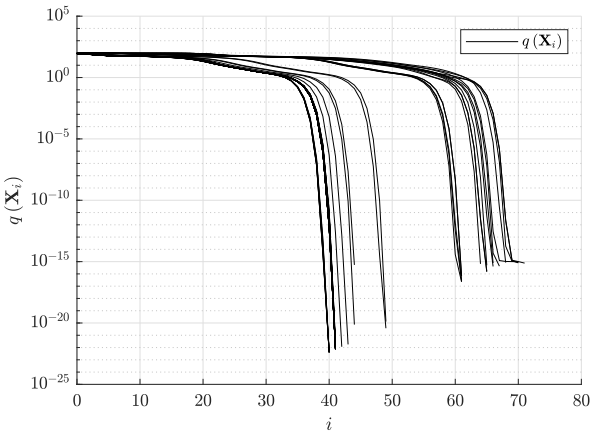


FIGURE 18. Convergence of scalar metric q obtained through the iterations i of the Schulz method for approximating the inverse of the Hessian matrix, while the initial guess was chosen as $\alpha\mathbf{H}^T$.

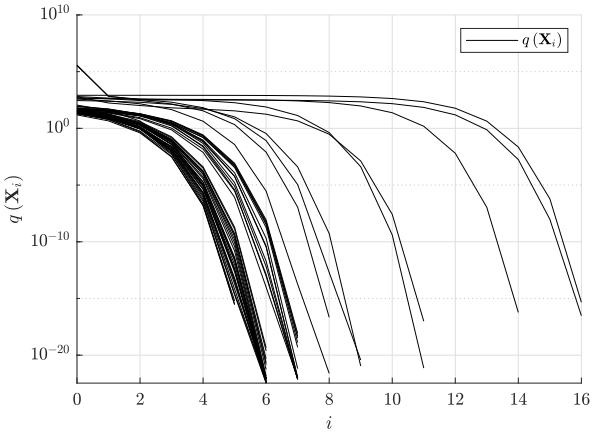
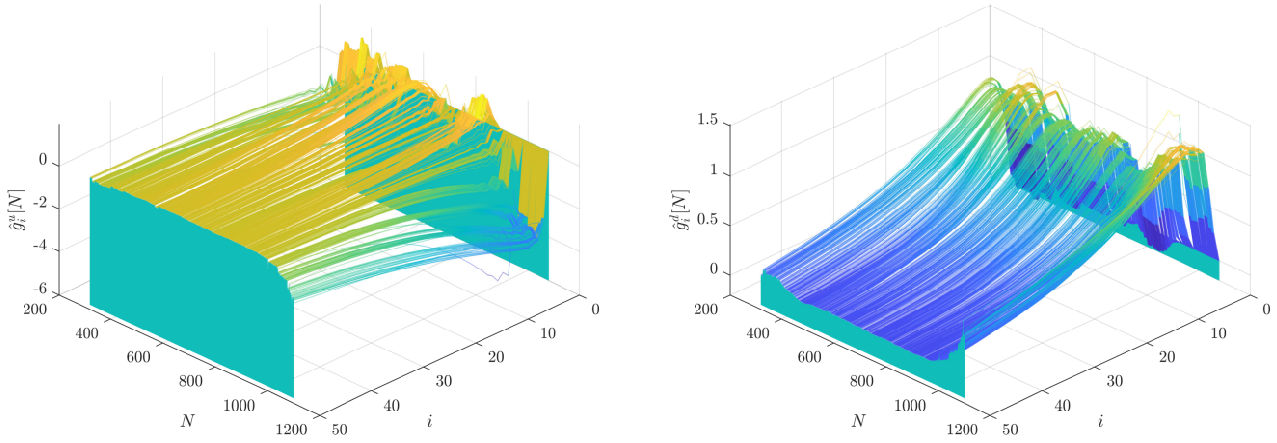


FIGURE 19. Convergence of scalar metric q obtained through the iterations i of the Schulz method for approximating the inverse of the Hessian matrix that was terminated after satisfying condition $q < 10^{-15}$, while the initial guess was chosen as $\alpha\mathbf{H}^{-1}[N]$.

The results presented above confirm that the proposed identification method can provide valid and accurate models for a whole spectrum of physiological variations in diabetic subjects.

Concerning the performance of the conjugate gradient method, the convergence of the cost function $f(\mathbf{x}_k)$ taken relative to the exact analytical solution $f(\mathbf{x}^*)$ is plotted in Figure 17. Notice that the actual convergence practically stops after reaching $M_u + M_d + 2 = 100$ iterations, while the approximation can be considered equal to the exact analytical solution, as this is typical characteristics of the conjugate gradient method applied to quadratic forms. Small fluctuations of the magnitude $< 10^{-10}$, which can be observed after the 100th iteration are caused just by rounding errors, hence this effect has a purely numerical nature.

Studying the Schulz method for approximating the inverse of the Hessian matrix while terminating iterative algorithm (70) after satisfying condition $q < 10^{-15}$, convergence of scalar metric q (83) is depicted in Figure 18 for the standard choice of the initial guess and in Figure 19 for the recursive choice of the initial guess. In Figure 19, relatively



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

FIGURE 20. **Online adaptive estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained in the terms of recursive least squares method applied to ARX model by processing the parameter-varying dataset.**

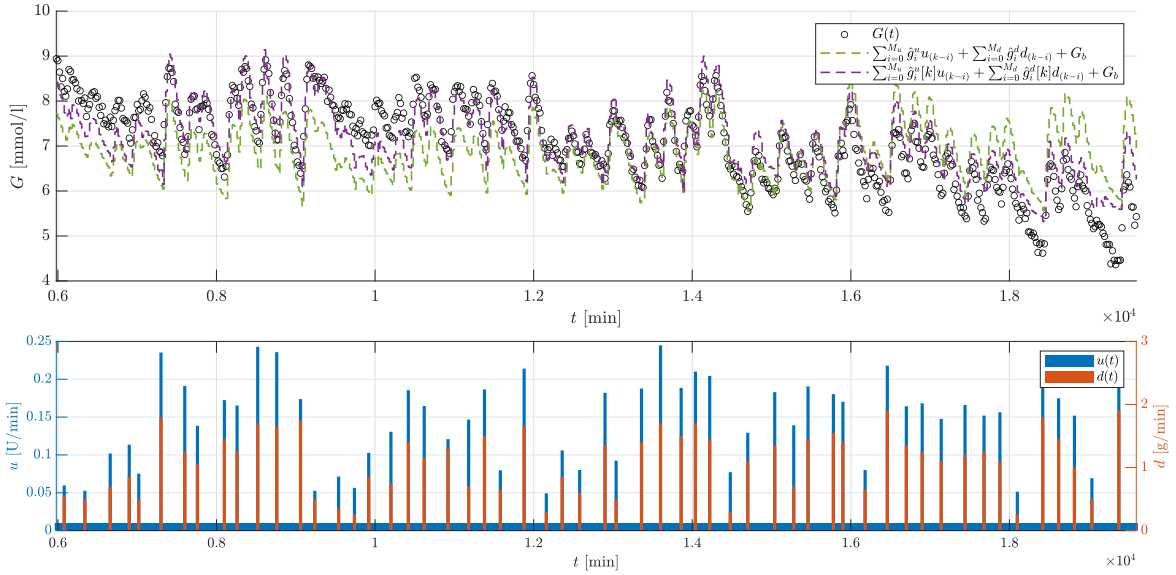


FIGURE 21. **Comparison of the output long-term prediction using the time-invariant model and the time-varying model.**

fast convergence is observed due to the recursive choice of the initial guess according to (78), while Figure 18 shows much slower convergence caused by the less effective choice of the initial guess according to (72).

The proposed correlation method will be compared with the recursive least squares algorithm (abbr. RLS) applied to minimize the prediction error of the two-input parameter-varying ARX, ARMAX and Box-Jenkins models.

The parameter-varying ARX model is defined in the z-domain as

$$A(z)[k]y_{(k)} = B_u(z)[k]u_{(k)} + B_d(z)[k]d_{(k)} + \epsilon_{(k)}\,, \qquad (85)$$

where the polynomials are of the degrees n_A for $A(z)$, n_B for $B_u(z)$ and $B_d(z)$.

Furthermore, the comparison will consider the parameter-varying ARMAX model defined as

$$A(z)[k]y_{(k)} = B_u(z)[k]u_{(k)} + B_d(z)[k]d_{(k)} + C(z)[k]\epsilon_{(k)}\,, \qquad (86)$$

where the polynomial $C(z)$ is of degree n_C.

We will also estimate the parameter-varying Box-Jenkins model

$$y_{(k)} = \frac{B_u(z)[k]}{A_u(z)[k]}u_{(k)} + \frac{B_d(z)[k]}{A_d(z)[k]}d_{(k)} + \frac{C(z)[k]}{D(z)[k]}\epsilon_{(k)}\,, \qquad (87)$$

where the polynomials are of degrees n_A for $A_u(z)$ and $A_d(z)$, n_B for $B_u(z)$ and $B_d(z)$, n_C for $C(z)$, n_D for $D(z)$.

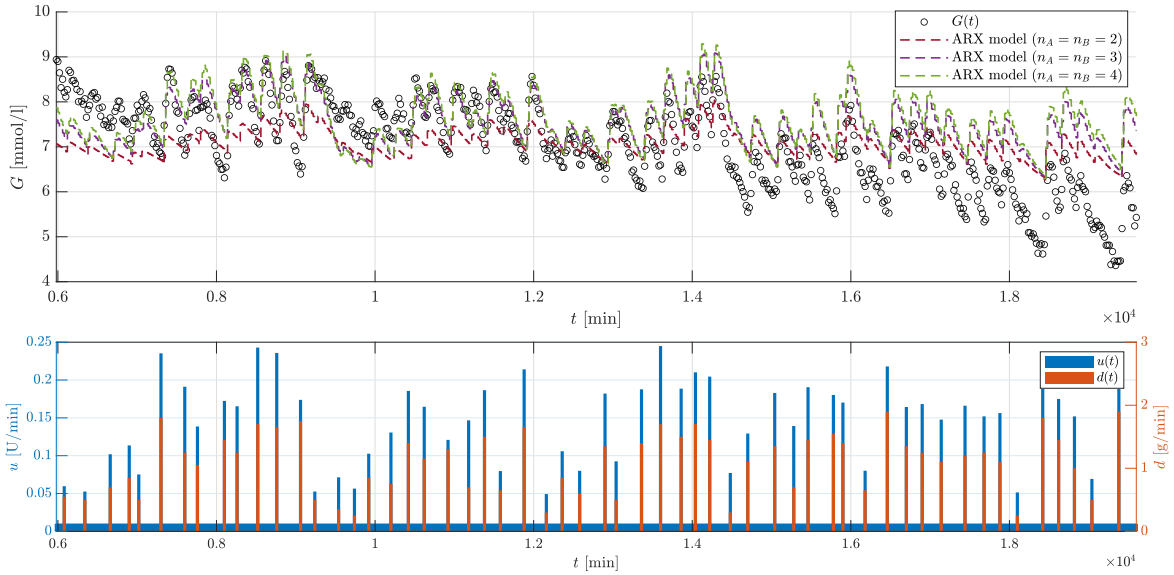


FIGURE 22. Long-term prediction of glycemia using the time-varying ARX model.

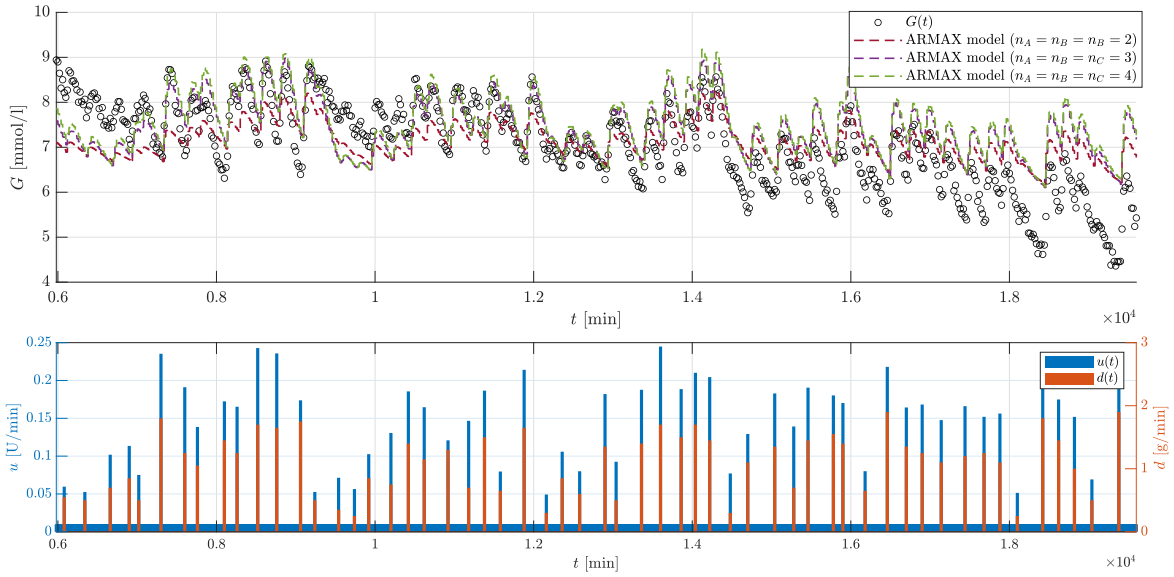


FIGURE 23. Long-term prediction of glycemia using the time-varying ARMAX model.

Equations of the RLS algorithm with the exponential forgetting will be adopted as [11]

$$Y[k] = \frac{P[k-1]h[k]}{\lambda + h^{\mathrm{T}}[k]P[k-1]h[k]}\,, \tag{88a}$$

$$P[k] = \frac{1}{\lambda}\left(I - Y[k]h^{\mathrm{T}}[k]\right)P[k-1]\,, \tag{88b}$$

$$\hat{\theta}[k] = \hat{\theta}[k-1] + Y[k]\left(y_{(k)} - h^{\mathrm{T}}[k]\hat{\theta}[k-1]\right), \tag{88c}$$

where y is the output measurement, Y is the correction vector, P is the covariance matrix of the parameter estimate, $\hat{\theta}$ is the vector of estimated parameters, λ is the forgetting factor, while the regression vector h is defined by the corresponding model structure (ARX, ARMAX, Box-Jenkins).

By application of the RLS algorithm (88) to estimate the ARX model (85) based on the input-output dataset from Figure 6 while considering $n_A = n_B = 3$, we obtained a time-varying model with evolution of the impulse response coefficients visualized in Figure 20. The figure also documents that this approach is not suitable for estimating models from diabetic data, as the obtained impulse responses are not compliant with the basic physiology of diabetes (particularly the insulin administration response is very defective compared to the standard shape in Figure 3), and hence its general validity is poor. This insufficient

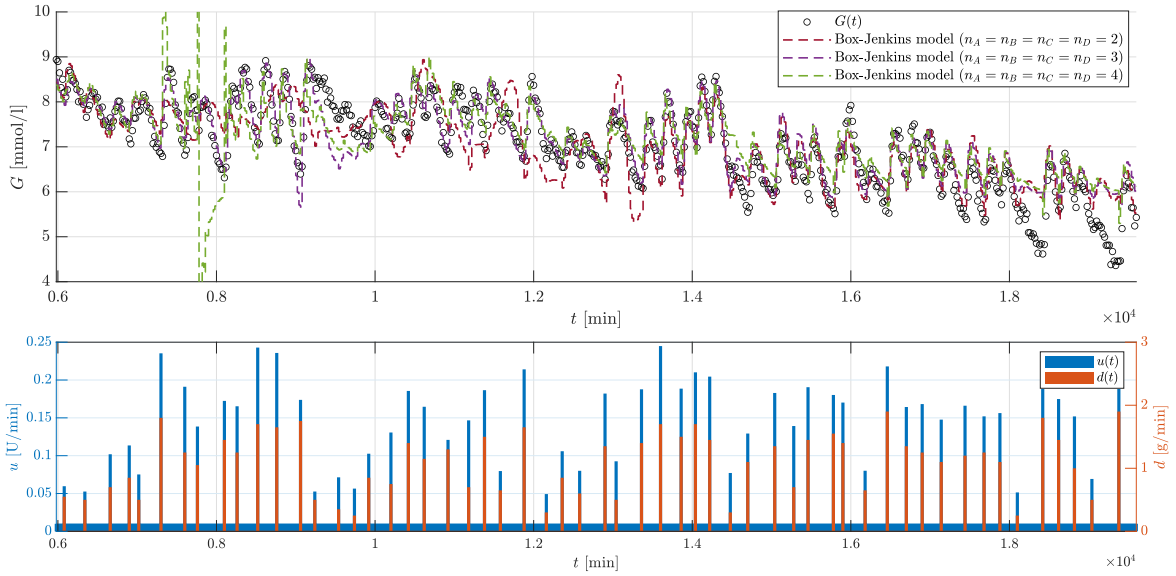
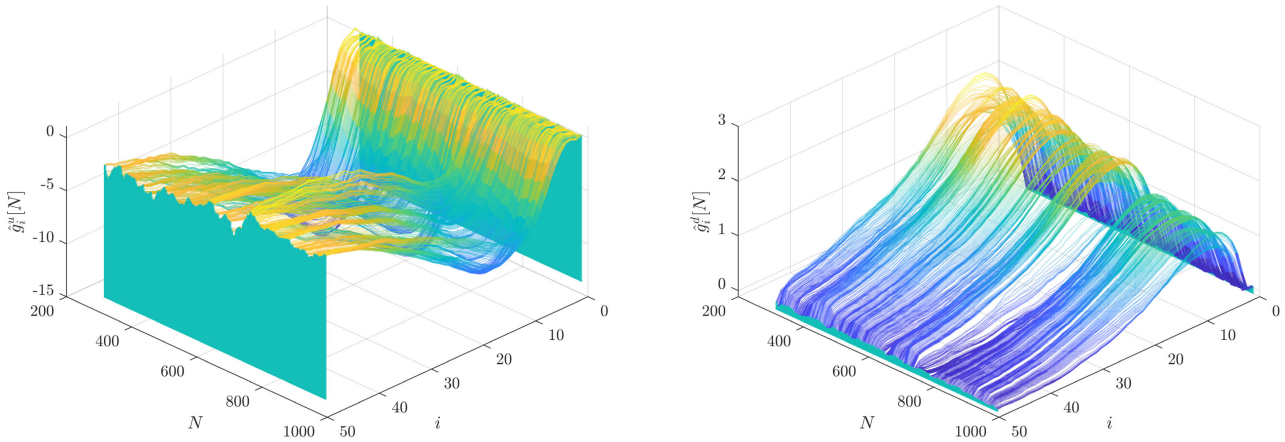


FIGURE 24. Long-term prediction of glycemia using the time-varying Box-Jenkins model.



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

FIGURE 25. Online adaptive estimates of the impulse response coefficients of the insulin administration effect submodel $\hat{g}^u[N]$ and the carbohydrate intake effect submodel $\hat{g}^d[N]$ as the functions of number of processed samples N, obtained by processing the parameter-varying dataset while considering $\lambda = 0.995$ and input uncertainties.

performance of the RLS algorithm can be explained by the poor excitation properties of the input signals $u_{(k)}, d_{(k)}$, which are impulsive in nature and are highly correlated as they occur simultaneously (see Figure 6).

Another evaluation of the results is aimed at long-term predicting (simulating) glycemia using the identified models. The output prediction of the nonparametric model can be calculated directly from the definition (1) for the time-invariant model and according to (2) in the case of time-varying model parameters. In Figure 21 there is the long-term prediction of glycemia performed on the parameter-varying dataset from Figure 6 assuming the offline identification strategy with parameters estimated as fixed, compared to the adaptation-based prediction of glycemia using the online estimation of time-varying impulse responses that were presented in Figure 12.

It can be concluded that the online identification yielded a time-varying model with better prediction performance than was obtained using the fixed offline-identified parameters, hence the model adapted better due to the intra-subject variability.

The predictions for the time-varying ARX model (85) of various orders are plotted in Figure 22, while the predictions for the time-varying ARMAX model (86) can be seen in Figure 23. It is apparent that the second-order ARX and ARMAX models performed insufficiently. The next comparison considers the time-varying Box-Jenkins model (87) with the corresponding predictions visualized

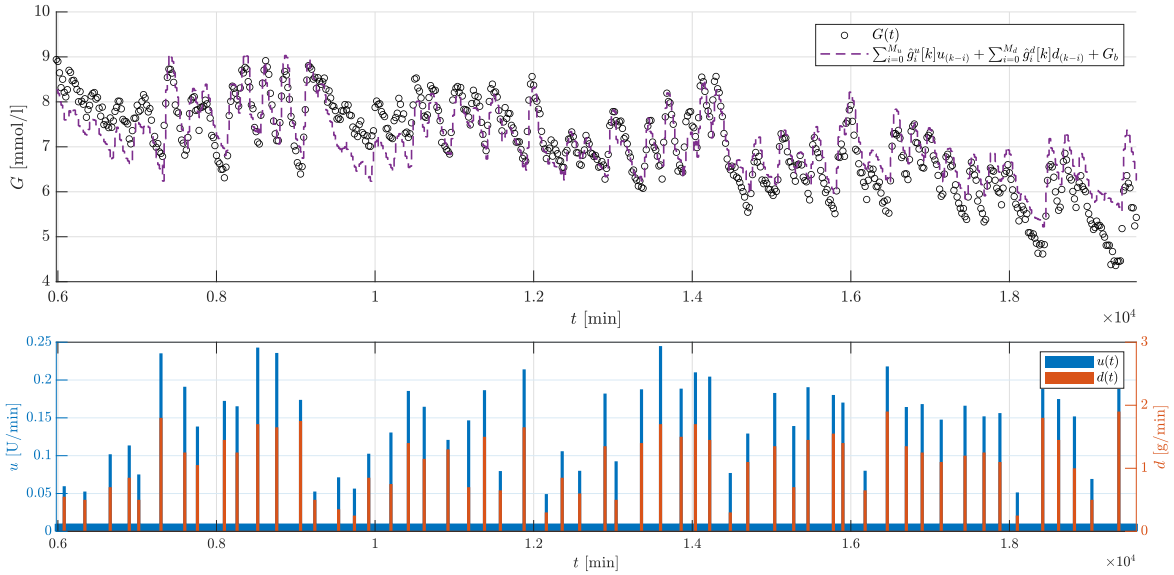


FIGURE 26. Long-term prediction of glycemia with input uncertainties using the time-varying model.

TABLE 2. Prediction performance metric.

setting	dataset	Q
$\alpha = 0, \beta = 0, \lambda = 1$	parameter-invariant	0.0144
$\alpha = 5 \times 10^2, \beta = 0, \lambda = 1$	parameter-invariant	0.0150
$\alpha = 5 \times 10^2, \beta = 1 \times 10^{-3}, \lambda = 1$	parameter-invariant	0.0152
$\alpha = 5 \times 10^2, \beta = 1 \times 10^{-3}$ (offline)	parameter-varying	0.6965
$\alpha = 5 \times 10^2, \beta = 1 \times 10^{-3}, \lambda = 0.995$	parameter-varying	0.1913
$\alpha = 5 \times 10^2, \beta = 1 \times 10^{-3}, \lambda = 0.995$	parameter-varying with input uncertainties	0.2109
$\alpha = 5 \times 10^2, \beta = 1 \times 10^{-3}, \lambda = 0.95$	parameter-varying	0.4851
ARX model $\lambda = 0.995, n_A = n_B = 2$	parameter-varying	0.6421
ARX model $\lambda = 0.995, n_A = n_B = 3$	parameter-varying	0.7838
ARX model $\lambda = 0.995, n_A = n_B = 4$	parameter-varying	0.9124
ARMAX model $\lambda = 0.995, n_A = n_B = n_C = 2$	parameter-varying	0.5471
ARMAX model $\lambda = 0.995, n_A = n_B = n_C = 3$	parameter-varying	0.6329
ARMAX model $\lambda = 0.995, n_A = n_B = n_C = 4$	parameter-varying	0.6961
Box-Jenkins model $\lambda = 0.995, n_A = n_B = n_C = n_D = 2$	parameter-varying	0.3689
Box-Jenkins model $\lambda = 0.995, n_A = n_B = n_C = n_D = 3$	parameter-varying	0.2064
Box-Jenkins model $\lambda = 0.995, n_A = n_B = n_C = n_D = 4$	parameter-varying	0.7088

in Figure 24. It can be claimed that the Box-Jenkins model generally provided more accurate predictions than the ARX and ARMAX models, while its performance is comparable to the nonparametric model obtained by the proposed correlation method. However, in Figure 24, there can be seen a tendency of the fourth-order Box-Jenkins model to instability, which is a highly undesired phenomenon.

In addition to standard experimentation with deterministic inputs and their effects, signals $u_{(k)}$, $d_{(k)}$ will be considered to be affected by random processes. The stochastic component of $d_{(k)}$ can be seen as the uncertainty in the magnitude of carbohydrate intake, which is typically not known precisely in real-life conditions. On the other hand, the stochastic component of $u_{(k)}$ represents the uncertain nature of insulin action, which can be affected by various factors, such as the site of administration.

To model the stochastic nature of inputs, signals $u_{(k)}$, $d_{(k)}$ entering the identification algorithm will be distorted as

$$u_{(k)} \cdot \chi_{u(k)}\,, \tag{89}$$

$$d_{(k)} \cdot \chi_{d(k)}\,, \tag{90}$$

where $\chi_{u(k)} \sim \mathcal{N}(1, 0.1)$ and $\chi_{d(k)} \sim \mathcal{N}(1, 0.1)$ follow the normal distributions.

The estimates of impulse responses obtained considering $\alpha = 5 \times 10^2$, $\beta = 1 \times 10^{-3}$, $\lambda = 0.995$ by processing the parameter-varying dataset from Figure 6 with uncertain inputs in terms of (89), (90) are documented in Figure 25. As expected, the impulse responses reacted to mismatch in input signals by significant perturbations, yet they still can be considered acceptable and usable in practice.

Then, in Figure 26 there is the corresponding long-term prediction of glycemia performed on the parameter-varying dataset from Figure 6 using the online estimation

of time-varying impulse responses that were presented in Figure 25. It can be concluded that the prediction accuracy is slightly deteriorated compared to Figure 21, yet satisfactory.

Evaluating metric (84) for the considered scenarios gives the results summarized in Table 2.

VIII. CONCLUSION

The paper presented an extension and a new theoretical framework for the correlation-based method to estimate impulse responses in the case of a two-input single-output nonparametric empirical model of type 1 diabetes. The augmented method now allows for effective online estimation and adaptation of the model parameters in real time by processing new available samples of the input-output signals and recursively updating correlation functions. To this end, the exponentially weighted estimate of the correlation function was introduced to implement the forgetting of older samples, while the recursive formula for updating this estimate was derived. The crucial Wiener-Hopf equation generalized for systems with two inputs was presented together with the corresponding linear regression system. Compared to the traditional correlation-based identification framework, we studied the statistical properties of the sample correlation function, which resulted in the estimation based on the generalized least squares method while minimizing the variance of the estimate. The cost function of the estimation problem was modified by considering two novel optimal regularization strategies. The first regularization term was designed to robustify the identification with respect to a priori known distribution of the true parameter vector. The second regularization term was designed to penalize the inter-sample change of the parameter estimate, i.e., drift of the parameter vector in time, and thus to suppress excessive and unlikely changes throughout the iterations. Following these regularization strategies, the overall robustness of the estimate was improved with respect to random disturbance factors, the presence of outliers, and the poor quality of the input-output data supplied.

One of the important findings to highlight is the recursive formula for updating the covariance matrix of the uncertainty of the correlation function estimate, but more importantly, the recursive formula for updating its inverse, which is actually essential to evaluate the generalized least squares method estimate.

Since we wanted to fully avoid calculating all the matrix inverses, the quadratic form corresponding to the cost function of the parameter estimation problem was, instead of solving the closed analytical form, numerically minimized using the conjugate gradient method that ensures convergence to the analytical solution in a finite number of iterations. Alternatively, the inverse Hessian matrix was proposed to be approximated using the Schulz method. Due to the derived criterion to ensure the method convergence using the Gershgorin circle theorem, the recursive choice of the initial approximation provided very fast convergence in practice.

The sequence of operations required to update the parameter estimate per sample was optimized with respect to the computational complexity as follows:

1) Performing recursive update of the sample correlation functions according to new input-output data instead of evaluating the full summation.
2) Performing recursive update of the inverse of the covariance matrix of the noise vector instead of its direct inversion.
3) Application of the conjugate gradient method or the Schulz method to obtain the parameter estimate without calculating the matrix inverse.

Because the number of iterations in the case of conjugate gradient algorithm is fixed and the complexity of other subroutines is independent of the character of arriving data, the number of arithmetic operations per sample is invariant and deterministic. This can be seen as an important factor when implementing periodic tasks in real-time operating systems. It can be claimed that the overall computational complexity depends only on the considered lengths of the impulse responses M_u, M_d (the number of estimated parameters) and the number of lags P, since these affect the dimensions of all concerned vectors and matrices. Due to the recursive formulation of the algorithm, the number of processed samples does not affect the computational complexity, which is a significant advantage in the case of real-time implementation. Concerning the target software implementation in an object-oriented programming language, all steps of the proposed estimation algorithm require a framework for basic operations with vectors and matrices, yet these exclude advanced and computationally demanding techniques such as matrix decompositions or the matrix inverse.

The results of simulation-based experiments confirmed that a properly tuned identification algorithm can produce satisfactory and valid estimates of the impulse functions even under demanding conditions caused by noisy glycemia measurements and highly correlated input signals typical of conventional insulin therapy. The proposed regularization strategies provided physiology-compliant estimates of impulse functions. Both alternative methods to solve the estimation problem showed good convergence properties and could replace the traditional explicit formula that involves the Hessian matrix inverse. Concerning the intra-subject variability, the online estimation strategy was able to adapt the identified impulse responses with respect to the time-varying parameters of the simulation model.

The presented comparison with the recursive least squares method for the estimation of the time-varying ARX, ARMAX and Box-Jenkins models showed that this conservative approach is not suitable for adaptive modeling of glycemia, despite the fact that it can provide sufficient prediction performance under specific circumstances.

An eventual implementation of the proposed online identification method and the adaptive model in some augmented

form of continuous glucose monitoring system would be appropriate and very likely to be the subject of future interest. This would benefit from an efficient combination of the glycemia monitoring device and the adaptive modeling approach while providing real-time updated information on the dynamic responses of glycemia to insulin administration and carbohydrate intake under the influence of time-varying physiological characteristics. These factors play pivotal roles in the design of techniques for maximizing the performance and safety of insulin treatment in diabetic subjects. Therefore, this form of model with adaptive capabilities is virtually mandatory for the implementation of an efficient artificial pancreas for automated continuous insulin dosing or for an advanced advisory system in the form of optimal model-based bolus calculator, which provides advice on the time and size of the insulin bolus to be administered. Emerging challenges in practical implementation of these solutions are of various nature, such as dealing with sensor failures, detecting mismatches in meal announcing, or managing the effects of random unmeasurable disturbances.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable suggestions that greatly improved the quality of this paper.

REFERENCES

[1] M. Dodek and E. Miklovicová, "Estimation of process noise variances from the measured output sequence with application to the empirical model of type 1 diabetes," *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104773.

[2] 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.

[3] S. Mehmood, I. Ahmad, H. Arif, U. Ammara, and A. Majeed, "Artificial pancreas control strategies used for type 1 diabetes control and treatment: A comprehensive analysis," *Appl. Syst. Innov.*, vol. 3, no. 3, p. 31, Jul. 2020.

[4] J. Tašic, M. Takács, and L. Kovács, "Control engineering methods for blood glucose levels regulation," *Acta Polytechnica Hungarica*, vol. 19, no. 7, pp. 127–152, 2022.

[5] S. J. Moon, I. Jung, and C.-Y. Park, "Current advances of artificial pancreas systems: A comprehensive review of the clinical evidence," *Diabetes Metabolism J.*, vol. 45, no. 6, pp. 813–839, Nov. 2021.

[6] 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.

[7] R. Sánchez-Peña and D. Chernavvsky, *Artificial Pancreas: Current Situation and Future Directions*. Cambridge, MA, USA: Academic Press, Apr. 2019.

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

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

[10] M. Dodek, E. Miklovicová, and M. Tárník, "Correlation method for identification of a nonparametric model of type 1 diabetes," *IEEE Access*, vol. 10, pp. 106369–106385, 2022.

[11] L. Ljung, *System Identification: Theory for the User* (Prentice-Hall Information and System Sciences Series). Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[12] 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.

[13] 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.

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

[15] D. Boiroux, A. K. Duun-Henriksen, S. Schmidt, K. Nørgaard, N. K. Poulsen, H. Madsen, and J. B. Jørgensen, "Adaptive control in an artificial pancreas for people with type 1 diabetes," *Control Eng. Pract.*, vol. 58, pp. 332–342, Jan. 2017.

[16] M. Eren-Oruklu, A. Cinar, D. K. Rollins, and L. Quinn, "Adaptive system identification for estimating future glucose concentrations and hypoglycemia alarms," *Automatica*, vol. 48, no. 8, pp. 1892–1897, Aug. 2012.

[17] G. C. Estrada, H. Kirchsteiger, L. del Re, and E. Renard, "Innovative approach for online prediction of blood glucose profile in type 1 diabetes patients," in *Proc. Amer. Control Conf.*, Jun. 2010, pp. 2015–2020.

[18] T. V. Herpe, M. Espinoza, B. Pluymers, I. Goethals, P. Wouters, G. V. D. Berghe, and B. D. Moor, "An adaptive input–output modeling approach for predicting the glycemia of critically ill patients," *Physiological Meas.*, vol. 27, no. 11, pp. 1057–1069, Sep. 2006.

[19] 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.*, 2008, pp. 2099–2104.

[20] H. Kirchsteiger, S. Pölzer, 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. 50th IEEE Conf. Decis. Control Eur. Control Conf.*, Dec. 2011, pp. 5176–5181.

[21] 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.

[22] 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, Jan. 2011.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] I. Rodríguez-Rodríguez, M. Campo-Valera, J.-V. Rodríguez, and W. L. Woo, "IoMT innovations in diabetes management: Predictive models using wearable data," *Exp. Syst. Appl.*, vol. 238, Mar. 2024, Art. no. 121994.

[28] 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 M. E. Wilinska, "Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes," *Physiological Meas.*, vol. 25, no. 4, pp. 905–920, Jul. 2004.

[29] C. Fabris and B. Kovatchev, *Glucose Monitoring Devices: Measuring Blood Glucose to Manage and Control Diabetes*. Amsterdam, The Netherlands: Elsevier, 2020.

[30] H. Kirchsteiger, J. Jørgensen, E. Renard, and L. del Re, *Prediction Methods for Blood Glucose Concentration: Design, Use and Evaluation* (Lecture Notes in Bioengineering). Berlin, Germany: Springer, 2016.

[31] G. Jenkins and D. Watts, *Spectral Analysis and Its Applications* (Holden-Day Series in Time Series Analysis and Digital Signal Processing). Oakland, CA, USA: Holden-Day, 1969.

[32] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[33] G. Box, G. Jenkins, G. Reinsel, and G. Ljung, *Time Series Analysis: Forecasting and Control* (Wiley Series in Probability and Statistics). Hoboken, NJ, USA: Wiley, 2015.
[34] T. Amemiya, *Advanced Econometrics*, 1st ed. Cambridge, MA, USA: Harvard Univ. Press, 1985.
[35] R. Davidson, C. Davidson, J. MacKinnon, and E. MacKinnon, *Econometric Theory and Methods*. Oxford, U.K.: Oxford Univ. Press, 2004.
[36] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2001.
[37] 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.
[38] 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, Aug. 2012.
[39] H. Pishro-Nik, *Introduction to Probability, Statistics, and Random Processes*. Kalamazoo, MI, USA: Kappa Research, LLC, 2014. [Online]. Available: https://www.probabilitycourse.com
[40] J. Sherman and W. J. Morrison, "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *Ann. Math. Statist.*, vol. 21, no. 1, pp. 124–127, Mar. 1950.
[41] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bur. Standards*, vol. 49, no. 6, pp. 409–435, Dec. 1952.
[42] R. Fletcher, "Function minimization by conjugate gradients," *Comput. J.*, vol. 7, no. 2, pp. 149–154, Feb. 1964.
[43] B. Poliak, *Introduction to Optimization* (Translations Series in Mathematics and Engineering). San Diego, CA, USA: Optimization Software, 1987.
[44] V. A. Tsachouridis, G. Giantamidis, S. Basagiannis, and K. Kouramas, "Formal analysis of the schulz matrix inversion algorithm: A paradigm towards computer aided verification of general matrix flow solvers," *Numer. Algebra, Control Optim.*, vol. 10, no. 2, pp. 177–206, 2020.
[45] S. Artidiello, A. Cordero, J. R. Torregrosa, and M. P. Vassileva, "Generalized inverses estimations by means of iterative methods with memory," *Mathematics*, vol. 8, no. 1, p. 2, Dec. 2019.
[46] G. Schulz, "Iterative berechung der reziproken matrix," *ZAMM-J. Appl. Math. Mech./Zeitschrift für Angew. Math. und Mech.*, vol. 13, no. 1, pp. 57–59, Jan. 1933.
[47] F. Riesz and B. Nagy, *Functional Analysis* (Dover Books on Mathematics). New York, NY, USA: Dover, 2012.
[48] M. Zima, "A theorem on the spectral radius of the sum of two operators and its application," *Bull. Austral. Math. Soc.*, vol. 48, no. 3, pp. 427–434, Dec. 1993.
[49] R. S. Varga, *Gersgorin and His Circles* (Springer Series in Computational Mathematics). Berlin, Germany: Springer, 2004.
[50] A. Ben-Israel, "A note on an iterative method for generalized inversion of matrices," *Math. Comput.*, vol. 20, no. 95, p. 439, Jul. 1966.
[51] 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, Oct. 2007.
[52] 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.
[53] 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.
[54] S. Schmidt and K. Nørgaard, "Bolus calculators," *J. Diabetes Sci. Technol.*, vol. 8, no. 5, pp. 1035–1041, Sep. 2014.

MARTIN DODEK was born in 1995. He received the M.Sc. degree in robotics and cybernetics and the Ph.D. degree from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology in Bratislava, in 2020 and 2023, respectively. Currently, he is a Postdoctoral Researcher interested in 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 Slovak University of Technology in Bratislava, in 1990 and 1997, respectively. Currently, she is with the Institute of Robotics and Cybernetics, Faculty of Electrical Engineering and Information Technology, STU in Bratislava. She is an Associate Professor. Her research interests include predictive control, adaptive control, system modeling, and identification.