

Approximate model-based state estimation in simplified Receding Horizon Control

Hemza Redjimi[†], József K. Tar^{†,‡,*}

[†]Doctoral School of Applied Informatics and Applied Mathematics,

[‡]Antal Bejczy Center for Intelligent Robotics, University Research and Innovation Center,

^{*}Institute of Applied Mathematics, John von Neumann Faculty of Informatics,

Óbuda University,

Budapest, Bécsi út 96/B, H-1034

Hungary

Received: January 10, 2021. Revised: January 29, 2021. Accepted: February 12, 2021.

Published: February 15, 2021.

Abstract- Multiple variable systems often have to be so controlled that in the lack of appropriate sensors no satisfactory information is available for the complete estimation of their state variables. Normally only their certain components are kept under observation and control, while the other ones evolve according to the consequences of the exerted control signal. In control-based treatment of patients suffering from "Type 1 Diabetes Mellitus (T1DM)", the only directly measured quantity is the subcutaneous glucose concentration in the blood controlled by a single control signal, the insulin ingress rate. The applied model may use several components in the state variable. The traditional "Receding Horizon Controller (RHC)" requires the estimation of the complete state variable for the calculation of the control signal. In this paper preliminary simulations are presented in which the operation of the RHC is studied in the control of two vertically connected, oscillating masspoints so coupled by springs that only the state of the upper one is observed and directly controlled. Instead sensor-based observations, the lower point's coordinate is calculated by the use of an available "rough" model. Preliminary calculations were made for a particular human glucose-insulin model, too. In the implementation of the RHC special simplifications were introduced. In our further work we wish to apply this method for investigating various T1DM treatment models.

Keywords- Gram-Schmidt Algorithm, Lagrange Multipliers, Model Predictive Control, Optimal Control, Receding Horizon Control, Reduced Gradient, State Estimation.

I. INTRODUCTION

IN process control tasks often multiple variable systems have to be so controlled that in the lack of appropriate sensors no satisfactory information is available for the complete estimation of their state variables. Normally only its certain components are kept under observation and control, while the other ones evolve according to the consequences of the exerted control signal without having any possibility for observing and/or keeping them at bay. Especially in life sciences, where the unobserved quantities can achieve lethal values, this practice is a serious problem. For instance, in control-based treatment of patients suffering from T1DM from the early "minimal model" developed by Bergman in 1986 [1] more and more sophisticated models were elaborated (e.g. [2, 3, 4, 5, 6, 7, 8, 9, 10]) and used for various control approaches (e.g. [11, 12, 13, 14, 15, 16, 17, 18]). For instance, in the paper [19] the glucose and insulin absorption "submodels" were described by the equations (1) - (4). They originated from the Cambridge model [20] and were inserted into the core model described by the equations (5) - (7).

$$\dot{D}_1(t) = -\frac{1}{\tau_D}D_1(t) + \frac{1000A_g}{M_wGV_G}Cd(t) , \quad (1)$$

$$\dot{D}_2(t) = -\frac{1}{\tau_D}D_2(t) + \frac{1}{\tau_D}D_1(t) , \quad (2)$$

$$\dot{S}_1(t) = -\frac{1}{\tau_S}S_1(t) + \frac{1}{V_I}u(t) , \quad (3)$$

$$\dot{S}_2(t) = -\frac{1}{\tau_S}S_2(t) + \frac{1}{\tau_S}S_1(t) , \quad (4)$$

$$\dot{G}(t) = -(p_1 + X(t))G(t) + p_1G_B + \frac{1}{\tau_D}D_2(t) , \quad (5)$$

$$\dot{X}(t) = -p_2 X(t) + p_3 (I(t) - I_B) , \quad (6)$$

$$\dot{I}(t) = -n(I(t) - I_B) + \frac{1}{\tau_S} S_2(t) . \quad (7)$$

The state variables in (1) - (7) are defined as follows. $D_1(t)$ and $D_2(t)$ mg/dL are the primary and secondary compartments belonging to glucose. $S_1(t)$ and $S_2(t)$ mU/L are the primary and secondary compartments belonging to insulin. Variable $G(t)$ mg/dL is the blood glucose (BG) concentration ("glycemia"), $I(t)$ mU/L is the blood insulin concentration and $X(t)$ 1/min is the insulin-excitabile tissue glucose uptake activity. The food intake done by the patient under treatment was handled as disturbance that was unknown for the controller. The only directly observable quantity was $G(t)$, and the only known quantity was the exerted control signal $u(t)$ mU/min. The other variables were calculated on the basis of the available **model parameters** of which only M_{wG} (the molecular weight of glucose), C (a conversion rate between various units), A_g (glucose utilization), G_B (basal glucose level), I_B (basal insulin level), BW (body weight as easily measurable quantity) were assumed to be exact, while the p_1, p_2, p_3 transfer rates, the n rate constant for insulin disappearance, the τ_D carbohydrate (CHO) to glucose absorption constant, V_I, V_G distribution volumes of insulin and glucose, and the τ_S insulin absorption constant were only approximately estimated. In the lack of measuring possibilities, these calculated values were used for setting the starting point of the new horizon of the Receding Horizon Controller outlined in the sequel.

As in a heuristic realization of the idea of the "Model Predictive Controllers (MPC)" (e.g. [21]), in the "Receding Horizon Control (RHC)" [22] a cost function constructed of the trajectory tracking error and control forces over a finite grid is minimized under the constraint that the dynamic model of the controlled system must be valid. Generally the problem can be formulated in (8) in which J means the cost function that normally expresses a "penalty cost" for the big tracking error depending on the coordinates \mathbf{x} , it may punish the application of too big control action denoted by \mathbf{u} , $\Phi(\mathbf{x}_F)$ can be an extra punishment as "terminal cost", function $\dot{\mathbf{x}} = \Psi(\mathbf{x}, \mathbf{u})$ describes the dynamic model of the controlled system, and Δt means the resolution of the discrete time-grid. While it is not compulsory to achieve exactly zero cost (some compromise between its often contradictory components can be achieved in this manner), the **constraint conditions** in the last line of (8) must be precisely met: this term expresses the "abilities" of the controlled system that must be taken into account in the realization of the cost minimization.

$$\begin{aligned} & \min_{\substack{\{\mathbf{x}_1, \dots, \mathbf{x}_F\} \\ \{\mathbf{u}_0, \dots, \mathbf{u}_{F-1}\}}} \left(\sum_{i=0}^{F-1} J(\mathbf{x}_i, \mathbf{u}_i) + \Phi(\mathbf{x}_F) \right) \\ & \text{subject to } \frac{\mathbf{x}_{i+1} - \mathbf{x}_i}{\Delta t} - \Psi(\mathbf{x}_i, \mathbf{u}_i) = \mathbf{0}. \end{aligned} \quad (8)$$

In general, for finding the local optimum under constraints a numerical realization of Lagrange "Reduced Gradient Algorithm" [23] must be implemented in the following manner. If the variables $\{\mathbf{x}_1, \dots, \mathbf{x}_F\}$ and $\{\mathbf{u}_0, \dots, \mathbf{u}_{F-1}\}$ can be simultaneously varied, in the appropriate grid points each component of the $\frac{\mathbf{x}_{i+1} - \mathbf{x}_i}{\Delta t} - \Psi(\mathbf{x}_i, \mathbf{u}_i)$ constraint term has to be associated with a Lagrange multiplier, and the following considerations can be done. If the local minimum of the differentiable multivariable function $J(x, u) : \mathbb{R}^n \mapsto \mathbb{R}$ is needed without constraints, instead of the actual point (x, u) , a neighboring point $(x, u) - \alpha \left(\frac{\partial J}{\partial(x, u)} \right)$ can be proved with a small positive number $0 < \alpha$. The function's value can decrease until the gradient becomes zero. If it is a priori known that the local minimum is 0, the process can be made faster by using the "Newton-Raphson Algorithm" in which α is so chosen that $-\alpha \|\partial J / \partial(x, u)\|^2 = J(x, u)$ (e.g. [24, 25]). If the optimization must be done under constraints expressed implicitly as $\{0 \equiv g_i(x, u) : \mathbb{R}^n \mapsto \mathbb{R}\} i = 1, \dots, K < n$, apart from particular cases, an $(n - K)$ dimensional hypersurface in \mathbb{R}^n is determined by the constraints, and the local optimum must be found over this hypersurface. The optimization can be started by finding a point on this hypersurface by using the Newton-Raphson Algorithm for e.g. $G(x, u) := \sum_{\ell} g_{\ell}^2(x, u)$, since $G(x, u) = 0$ if and only if each $g_{\ell}(x, u) = 0$. In the computations the formula given in (9) can be used

$$(x, u)_{s+1} = (x, u)_s - a_1 \times \frac{G((x, u)_s)}{\left(\frac{\partial G}{\partial(x, u)} \right)^T \Big|_{(x, u)_s} \frac{\partial G}{\partial(x, u)} \Big|_{(x, u)_s}} \frac{\partial G}{\partial(x, u)} \Big|_{(x, u)_s} \quad (9)$$

in which $0 < a_1 < 1$. Following that, to avoid leaving the constraint-defined hypersurface while decreasing the cost function a small step can be done along the direction of the "reduced gradient" as $a_2 (-\partial J(x, u) / \partial(x, u) + \sum_{\ell} \lambda_{\ell} \partial g_{\ell} / \partial(x, u))$ in which $0 < a_2 < 1$, and the $\{\lambda_{\ell}\}$ Lagrange multipliers must be so chosen that the displacement must be orthogonal to each constraint gradient $\partial g_i(x, u) / \partial(x, u)$, because in the first order approximation then the constraint functions stagnate at 0. By using the 0 scalar product for expressing orthogonality, according to Lagrange, a linear set of equations is obtained and has to be solved as

$$\begin{aligned} \forall k : & - \left(\frac{\partial g_k}{\partial(x, u)} \right)^T \frac{\partial J}{\partial(x, u)} + \\ & + \sum_{\ell} \left(\frac{\partial g_k}{\partial(x, u)} \right)^T \frac{\partial g_{\ell}}{\partial(x, u)} \lambda_{\ell} = 0 . \end{aligned} \quad (10)$$

This process can be repeated until the reduced gradient achieves zero. In this point a local optimum allowed by the constraints has been found. (Sometimes the use of (9) may be necessary to correct the accumulated higher order errors to precisely stick on the "constraint hypersurface".) By our days this approach became a practical problem solving tool in financial and technical

problems as it is implemented in the "Solver" package of MS EXCEL (e.g. [26, 27]).

In many cases the Lagrange multipliers have important physical meaning, therefore it is often necessary to compute them (e.g. [28, 29]). Via the introduction of the "Auxiliary Function" in (11)

$$\Phi(x, u, \lambda) := -J(x, u) + \sum_{\ell} \lambda_{\ell} g_{\ell}(x, u) \quad (11)$$

it can be observed that the above detailed algorithm stops at the x, u, λ values at which $\partial\Phi/\partial x = 0$, $\partial\Phi/\partial u = 0$, and $\partial\Phi/\partial\lambda = 0$. The first requirement means that at this point the reduced gradient is 0, the second one guarantees that the appropriate point is located on the constraint surface.

Due to the special structure of the constraints in (8), for a fine enough grid resolution these partial derivatives can be interpreted as *canonical equations of motion* for the *canonical variable pairs* x and λ using an *artificial Hamiltonian* $H(x, u, \lambda) = -J(x, u) - \lambda^T \Psi(x, u)$ as $\partial H/\partial x = \dot{\lambda}$, $\partial H/\partial \lambda = -\dot{x}$, and $\partial H/\partial u = 0$ for the optimized motion over the grid. This formal analogy allows the utilization of basic results in Classical Mechanics as e.g. the properties of the Symplectic Geometry, Liouville's theorem [30] on the conservation of the volume of the phase space determined by the cells $\Delta x, \Delta \lambda$ from which general stability consequences can be drawn. For instance, in the case of Linear Time-invariant (LTI) system models and quadratic cost functions of the form $J = x^T P x + u^T R u$ with positive definite symmetric P and R matrices the Linear Quadratic Regulator's [31] equations can be obtained. In this special case from the equation of motion for λ , with a complementary trick, the Riccati equations [32] can be obtained that, though they are quadratic first order equations, can be solved by using the solution of second order linear equations. In commercial software products as e.g. MATLAB efficient solutions can utilize these possibilities for LQR. In Classical Thermodynamics the auxiliary function and the Lagrange multipliers give the basis for introducing the Chemical Potentials via using the Legendre transformation (e.g. [28]). The above "cultural reasons" mean strong motivations for engineers to proceed along these "tracks" (e.g. [33, 34, 35]).

Since the above numerical algorithm has considerable computational needs, it is an interesting question whether it can be avoided by using simpler techniques. For more general cases simplifications were suggested in [36] by abandoning the use of the auxiliary function. The basic idea was based on the observation that for solving the set of linear equations in (10) for the Lagrange multipliers can be realized by the use of the simple Gram-Schmidt algorithm [37, 38] (originally was invented by Laplace [39]) in the following manner. Let the component of vector a subtracted from vector b in the form $\tilde{b} := b + \mu a$ so that \tilde{b} must be orthogonal to a . Since orthogonality means zero scalar product solving the simple equation $0 = a^T \tilde{b} + \mu a^T a$ yields the appropriate real number μ . If the reduction of the gradient is important,

the appropriate vectors can be arranged as the columns of a matrix in (12).

$$\left[\begin{array}{c|c|c|c} \frac{\partial g_1}{\partial(x,u)} & \dots & \frac{\partial g_K}{\partial(x,u)} & -\frac{\partial J}{\partial(x,u)} \end{array} \right] \quad (12)$$

At first remove the component of the 1st column from the others. Then select the so obtained 2nd column and subtract its component from the remaining ones, etc. Finally, in the last column of the result the reduced gradient will be present. Further simplification can be done if we take into account that $\forall i g_i(x, u) = 0$ if and only if $G(x, u) = 0$, therefore it is enough to make a single step of orthogonalization for the two columns matrix $\left[\begin{array}{c|c} \frac{\partial G}{\partial(x,u)} & -\frac{\partial J}{\partial(x,u)} \end{array} \right]$. In [36] it was found the even in the case of a very simple paradigm quite considerable computation effort was spared in this manner. In the next section further simplifications are recommended for the purposes of developing RHC controllers using nonspecial dynamic models and cost functions.

II. FURTHER SIMPLIFICATIONS FOR RHC CONTROLLERS

It is a plausible observation that if we wish to release ourselves from the formalism using the Lagrange multipliers, in the case of a first order system it is just enough to execute the optimization according to the control signals over the horizon \mathbf{u}_i for $i = 0, 1, \dots, F - 1$, because in the possession of the system model the "constraint term" can be explicitly utilized as follows: \mathbf{x}_0 is given as the *initial condition*. The principle of causality can be expressed by the use of forward differences: $\mathbf{x}_1 = \mathbf{x}_0 + \Delta t \Psi(\mathbf{x}_0, \mathbf{u}_0)$, $\mathbf{x}_2 = \mathbf{x}_1 + \Delta t \Psi(\mathbf{x}_1, \mathbf{u}_1)$, etc. In this case over the horizon we have to minimize only $J(\mathbf{u}_0, \dots, \mathbf{u}_{F-1})$ with the simple gradient descent method without constraints. Neither the Newton-Raphson phase depicted in (9), nor moving along the reduced gradient by either solving (10) or making the reduction according to (12) has to be done. This evidently means quite considerable reduction in the necessary burden of computations.

This possibility is illustrated by the example of the vertical motion of two coupled mass-points of masses m_1 and m_2 in a gravitational field of acceleration g . The mass-point m_1 is hanging on a spring of stiffness k_1 and "zero force length" L_{01} . Its motion can be controlled by the vertical force F_1 . As "parasite dynamics", mass-point m_2 is connected to m_1 by a spring of stiffness k_2 and "zero force length" L_{20} . The "exact equations of motion" of this system are given in

$$\begin{aligned} m_1 \ddot{q}_1 &= m_1 g - k_1(q_1 - L_{01}) + k_2(q_2 - q_1 - L_{20}) - \\ &\quad - b_1 \dot{q}_1 + F_1 \\ m_2 \ddot{q}_2 &= m_2 g - k_2(q_2 - q_1 - L_{20}) - b_2 \dot{q}_2 \end{aligned} \quad (13)$$

where q_1 and q_2 are the appropriate vertical coordinates (increasing downwards), and b_1 and b_2 are viscous damping coefficients. It is assumed that q_1, \dot{q}_1 and \ddot{q}_1 can be directly observed, but no any measurable information is available on q_2, \dot{q}_2 and \ddot{q}_2 . Furthermore, it is assumed

that none of the exact parameters of the model $m_1, m_2, k_1, k_2, L_{01}, L_{02}, b_1, b_2,$ and g are available. Instead of them their approximations are available as $\tilde{m}_1, \tilde{m}_2, \tilde{k}_1, \tilde{k}_2, \tilde{L}_{01}, \tilde{L}_{02}, \tilde{b}_1, \tilde{b}_2,$ and \tilde{g} (Table 1). For the not observable variable its "computed estimation" is used as $\tilde{q}_2, \dot{\tilde{q}}_2$ and $\ddot{\tilde{q}}_2$, and when the estimation happens over the horizon during the optimization, the "optimized data over the horizon" $\check{q}_1, \dot{\check{q}}_1$ and $\ddot{\check{q}}_1$ are substituted in the approximate model equation of motion

$$\begin{aligned} \tilde{m}_1 \ddot{\check{q}}_1 &= \tilde{m}_1 \tilde{g} - \tilde{k}_1 (\check{q}_1 - \tilde{L}_{01}) + \tilde{k}_2 (\tilde{q}_2 - \check{q}_1 - \tilde{L}_{02}) - \\ &\quad - \tilde{b}_1 \dot{\check{q}}_1 + F_1 \\ \tilde{m}_2 \ddot{\tilde{q}}_2 &= \tilde{m}_2 \tilde{g} - \tilde{k}_2 (\tilde{q}_2 - \check{q}_1 - \tilde{L}_{02}) - \tilde{b}_2 \dot{\tilde{q}}_2, \end{aligned} \quad (14)$$

in which the data for \check{q}_1 , and \tilde{q}_2 are calculated over the horizon. The data over a horizon consists of the values $\{\check{q}_1[i], \dot{\check{q}}_1[i], \ddot{\check{q}}_1[i], \tilde{q}_2[i], \dot{\tilde{q}}_2[i], \ddot{\tilde{q}}_2[i], F_1[i] | i \in \{1, 2, \dots, H_L\}\}$ where $4 \leq H_L \in \mathbb{N}$ denotes the horizon length.

Table 1: The model parameters

Parameter	Exact Value	Approx. Value
Mass 1	$m_1 = 1.0 \text{ kg}$	$\tilde{m}_1 = 1.5 \text{ kg}$
Mass 2	$m_2 = 2.0 \text{ kg}$	$\tilde{m}_2 = 2.5 \text{ kg}$
Spring stiffness 1	$k_1 = 15.0 \text{ N} \cdot \text{m}^{-1}$	$\tilde{k}_1 = 11.0 \text{ N} \cdot \text{m}^{-1}$
Spring stiffness 2	$k_2 = 12.0 \text{ N} \cdot \text{m}^{-1}$	$\tilde{k}_2 = 16.0 \text{ N} \cdot \text{m}^{-1}$
Spring's length 1	$L_{01} = 1.0 \text{ m}$	$\tilde{L}_{01} = 1.2 \text{ m}$
Spring's length 2	$L_{02} = 1.5 \text{ m}$	$\tilde{L}_{02} = 1.8 \text{ m}$
Damping coeff. 1	$b_1 = 0.8 \text{ N} \cdot \text{s} \cdot \text{m}^{-1}$	$\tilde{b}_1 = 0.3 \text{ N} \cdot \text{s} \cdot \text{m}^{-1}$
Damping coeff. 2	$b_2 = 0.5 \text{ N} \cdot \text{s} \cdot \text{m}^{-1}$	$\tilde{b}_2 = 0.2 \text{ N} \cdot \text{s} \cdot \text{m}^{-1}$
Grav. accel.	$g = 9.81 \text{ m} \cdot \text{s}^{-2}$	$\tilde{g} = 10.0 \text{ m} \cdot \text{s}^{-2}$

Instead using Lagrange multipliers within the horizon Euler integration happens in the "forward form" as follows:

$$\begin{aligned} \check{q}_1[i+1] &= \check{q}_1[i] + \Delta t \dot{\check{q}}_1[i], \\ \tilde{q}_2[i+1] &= \tilde{q}_2[i] + \Delta t \dot{\tilde{q}}_2[i], \\ \dot{\check{q}}_1[i+1] &= \dot{\check{q}}_1[i] + \Delta t \ddot{\check{q}}_1[i], \\ \dot{\tilde{q}}_2[i+1] &= \dot{\tilde{q}}_2[i] + \Delta t \ddot{\tilde{q}}_2[i], \\ [\ddot{\check{q}}_1[i], \ddot{\tilde{q}}_2[i]] &= \widetilde{Mod}(\check{q}_1[i], \tilde{q}_2[i], \dot{\check{q}}_1[i], \dot{\tilde{q}}_2[i], F_1[i]), \end{aligned} \quad (15)$$

where the "Approximate Model Function" \widetilde{Mod} is defined in (14). The horizon is related to the "Exact Model" given in (13) by the "initial conditions" within the horizon: $\check{q}_1[1] = q_1(t), \dot{\check{q}}_1[1] = \dot{q}_1(t)$ and $\tilde{q}_2[1] = \tilde{q}_2(t), \dot{\tilde{q}}_2[1] = \dot{\tilde{q}}_2(t)$ in which the actual time t corresponds to the first grid point of the horizon. According to (15), the optimization happens for the freely selectable $F_1[1], F_1[2], \dots, F_1[H_L - 2]$ values, since the effect $F_1[i]$ appears in $\check{q}_1[i+2]$ in the case of a second order system, and the cost function in the grid point i is defined in (16)

$$J[i] = \begin{cases} C[i] \left(\frac{|q_1^N[i] - \check{q}_1[i]|}{\Delta q[i]} \right)^{P[i]} & \text{if } \frac{|q_1^N[i] - \check{q}_1[i]|}{\Delta q[i]} > 1 \\ C[i] \left(\frac{|q_1^N[i] - \check{q}_1[i]|}{\Delta q[i]} \right)^2 & \text{otherwise.} \end{cases} \quad (16)$$

where $i \in \{3, 4, \dots, H_L\}$, $q_1^N[i]$ denotes the nominal trajectory to be tracked in the grid point i , $C[i] > 0$ are weighting factors, $P[i] > 2$ determine how rapidly increases the penalty in the $\frac{|q_1^N[i] - \check{q}_1[i]|}{\Delta q[i]} > 1$ region. Since this power form produces a rapid increase in this region but varies very slowly if $\frac{|q_1^N[i] - \check{q}_1[i]|}{\Delta q[i]} \leq 1$, for this small values the quadratic cost produces nicely decreasing gradient. Consequently, the parameters $\Delta q[i]$ influence the "precision of tracking": the error of order of magnitude $\Delta q[i]$ is well tolerated but the higher values are drastically "punished". In a given grid point $C[i]$ and $P[i]$ have "redundant" effects, but their simultaneous use is not irrelevant. The total cost to be minimized is $J(F) = \sum_{i=3}^{H_L} J[i]$ in which $\mathbb{R}^{H_L-2} \ni [F_1, \dots, F_{H_L-2}] = [F_1[1], \dots, F_1[H_L - 2]]$. Because the "constraint terms" were eliminated by explicit calculations the simple gradient descent method can be applied for minimizing $J(F)$. The minimization was initiated by using $\forall i F_1[i] = 0$ starting values. Following the minimization the "optimized" $F_1(t) = F_1[1]$ is used in the Euler integration based on (13) as $q_1(t + \Delta t) \approx q_1(t) + \Delta t \dot{q}_1(t)$, $\dot{q}_1(t + \Delta t) \approx \dot{q}_1(t) + \Delta t \ddot{q}_1(t)$, and $q_2(t + \Delta t) \approx q_2(t) + \Delta t \dot{q}_2(t)$, $\dot{q}_2(t + \Delta t) \approx \dot{q}_2(t) + \Delta t \ddot{q}_2(t)$ for the calculation of the actual state variables. The values $\tilde{q}_2[2], \dot{\tilde{q}}_2[2], q_1(t + \Delta t), \dot{q}_1(t + \Delta t)$ are selected and put to the first grid point of the next horizon.

III. SIMULATION RESULTS FOR COUPLED MASS POINTS

For the simulations no any special commercial software product was used. The Julia language [?] was applied under Manjaro Linux on a DELL inspiron 15R computer. A simple own-made sequential code was developed. In the simulations $\Delta t = 10^{-3} \text{ s}$ discrete time-resolution, $H_L = 4$ horizon length, $\forall i C[i] = 1$ weight factors, $\forall i P[i] = 3$ cost powers, $\forall i \Delta q[i] = 10^{-3} \text{ m}$ tracking precision parameters, $\alpha_{max} = 10^{-2}$, $\alpha_{min} = 10^{-5}$ maximal and minimal speed parameter in the gradient descent algorithm was used with maximally allowed 200 consecutive steps. Within each optimization cycle the starting value was $\alpha = \alpha_{max}$. This value was kept until the norm $\|\nabla J(F)\|$ was decreasing. If in the next step this value was increased, the algorithm went back to the previous point and tried the next point with a smaller value $\alpha = \alpha_{previous}/1.2$ to make finer estimation. The initial state of the whole computation approximately was the equilibrium state of the free system (i.e. when $F_1 \equiv 0$). The results are given in Figs. 1–3. Figure 1 reveals that the effects of the initial velocity discrepancies are slowly relaxed in tracking the nominal trajectory $q_1^N(t)$. For

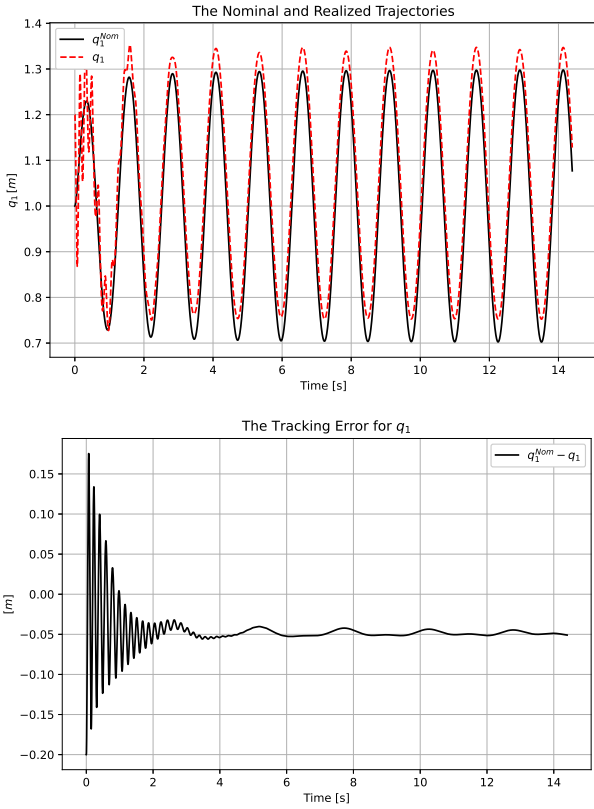


Fig. 1: Trajectory tracking and trajectory tracking error for $q_1^N(t)$

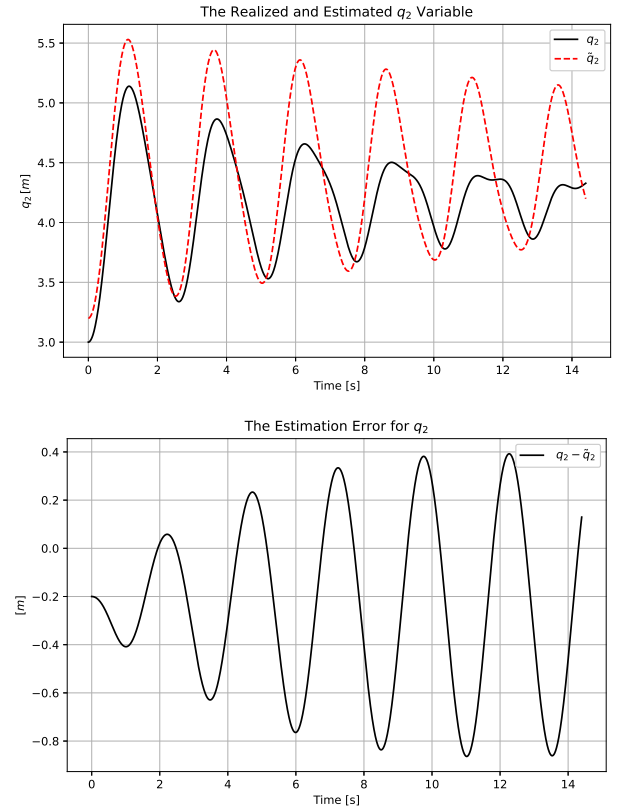


Fig. 2: The exact and the estimated value and the estimation error of $q_2(t)$

the success of this strategy the viscous damping that is present in the system plays an important role.

According to Fig. 2 it can be stated that the estimation error for the state variable component q_2 is significant but at least it is kept at bay. Also, it can be stated that as the initial oscillations that were caused by the initial state discrepancies decrease, the necessary control force is "tamed", too (Fig. 3).

To check the reliability of these results the simulation was repeated for "exact estimated model parameters". As it was expected, in the possession of the exact model parameters, Figs. 4 and 5 reveal definite relaxation of the initial tracking error. According to Fig. 6 the fluctuation of the control force is much smaller than in the case of very erroneous model parameter estimation.

It is important to note that the present considerations give answer only for the time-dependence of the estimation $\tilde{q}_2(t)$ by computing $\tilde{q}_2(t)$ from some approximate model. For this time-dependence the initial conditions $\tilde{q}_2(t_0)$ and $\dot{\tilde{q}}_2(t_0)$ give considerable contributions. For the estimation of the initial conditions complementary principles are needed that have to be relevant in the subject area of the phenomena under consideration.

Since our main aim is later making extensive investigations for T1DM treatment by the use of the here outlined method, in the sequel we give preliminary simulations for the same mathematical model that was treated by the MS EXCEL, VBA, Solver combination in [19].

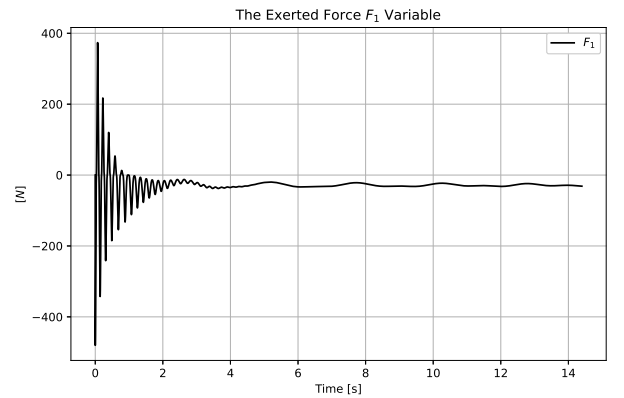


Fig. 3: The control force

IV. SIMULATIONS FOR A PARTICULAR T1DM MODEL

In [19] a T1DM case was investigated with the definite aim of using the resources of the MS EXCEL's Solver package. On this reason (8) was considered, i.e. the dynamic model was treated by the use of Lagrange multipliers as a constraint. To exemplify the here suggested method's applicability certain computations were made for this model, too. For the equations (1)–(7) the "exact parameters" are given in Table 2.

Utilizing the fact that in this approach *the limitation to the use of quadratic cost functions* was not necessary,

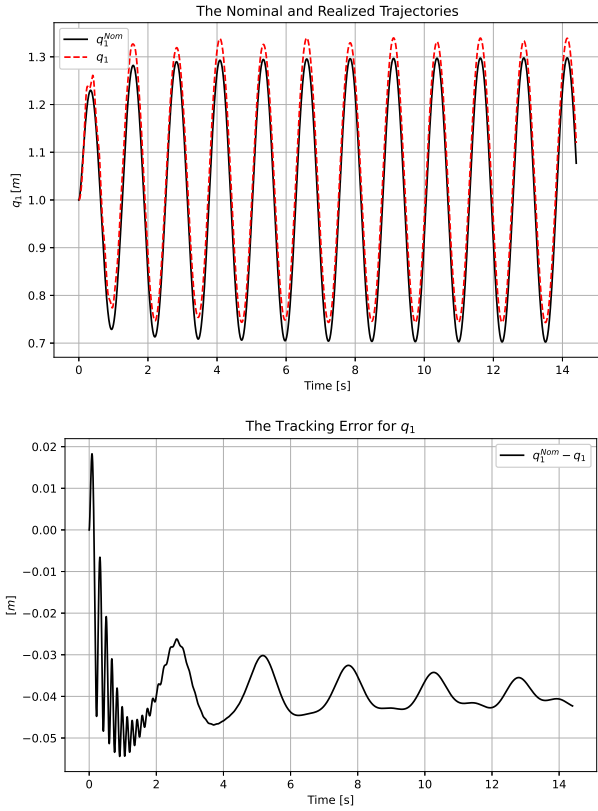


Fig. 4: Trajectory tracking and tracking error for $q_1^N(t)$ for exact estimated model parameters

in each relevant grid point $i = 2, \dots, F$ contributions penalizing the tracking errors and that punishing the use of too large insulin ingress rates were introduced as

$$\Psi = \sum_i (\Phi_G[i] + \Phi_u[i]),$$

$$\Phi_G[i] = C_G \begin{cases} \left(\frac{|G^N - G_o[i]|}{\Delta_G} \right)^{P_G} & \text{if } |G^N - G_o[i]| \geq \Delta_G \\ \left(\frac{|G^N - G_o[i]|}{\Delta_G} \right)^2 & \text{otherwise,} \end{cases}$$

$$\Phi_u[i] = C_u \begin{cases} \left(\frac{|G^N - G_o[i]|}{\Delta_G} \right)^{P_u} & \text{if } |u[i]| \geq \Delta_u \\ \left(\frac{|u[i]|}{\Delta_u} \right)^4 & \text{otherwise,} \end{cases}$$

that, with the used $P_G = P_u = 2.9$, means strongly increasing penalty for the "too big" tracking error and control components. Keeping the quadratic structure for the low tracking errors caused nice trajectory tracking, while the use of the 4th power for small insulin ingress rates practically did not mean any significant penalization. In the cost functions the $C_G = C_u = 1$ relative weights were used. The appropriate "ranges" were determined by the parameters $\Delta_G = 0.5$ mg/dL, and $\Delta_u = 30$ mU/min. The time-resolution of the grid was $\Delta t = 1$ min, the horizon length was 10 steps. The initial state corresponded to a *stable stationary state belonging to a low positive constant insulin ingress rate that stabilizes a given glycemia level*.

For practical investigations various "scenarios" can

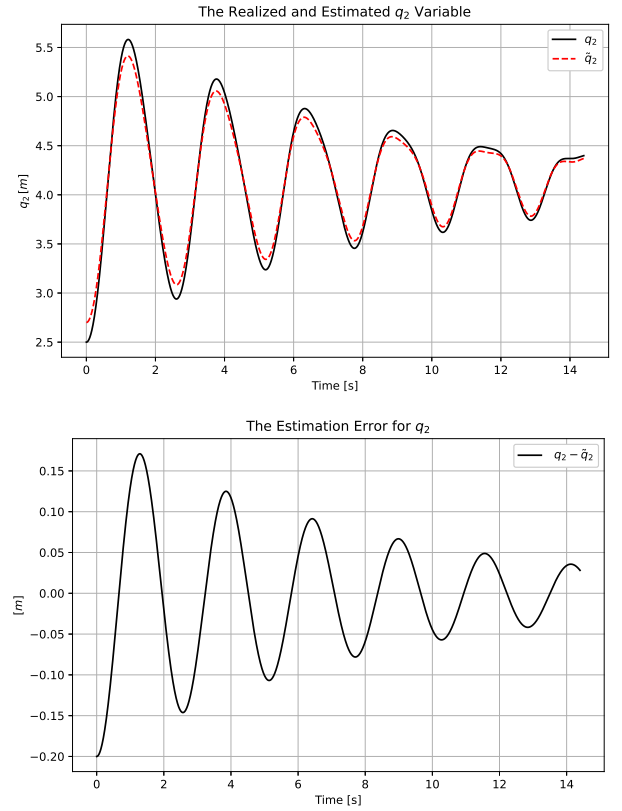


Fig. 5: Trajectory tracking and tracking error for $q_2^N(t)$ for exact estimated model parameters

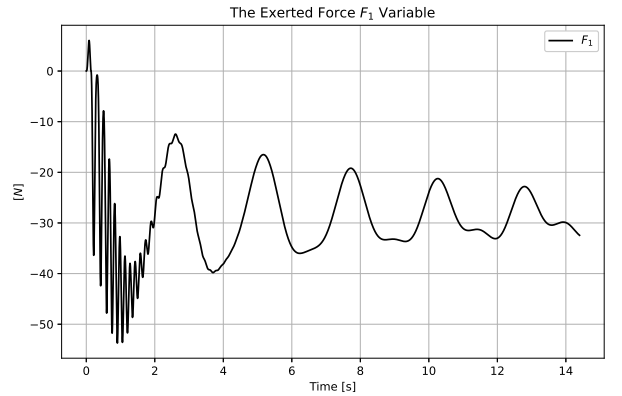


Fig. 6: The control force for exact approximate parameters

be invented concerning the "speed" of the "glucose subsystem" and the "insulin subsystem". It is an interesting question what happens if in the available approximate model the dynamics of the insulin subsystem is slower, while the dynamics of the COH intake subsystem is faster than the appropriate subsystems of the exact model. It is very interesting, too, how the RHC controller works when the approximate model's insulin subsystem is faster, and its COH intake subsystem is slower than that of the exact model. For investigating these situations two scenarios were considered belonging to "**Scenario A**" and "**Scenario B**", respectively. The

Table 2: The assumed "exact" model parameters used for the equations (1)–(7) in [19].

Notation	Value
G_B	$110.0 \frac{\text{mg}}{\text{dL}}$
I_B	$1.5 \frac{\text{mu}}{\text{L}}$
p_1	$2.8 \times 10^{-2} \frac{1}{\text{min}}$
p_2	$2.5 \times 10^{-2} \frac{1}{\text{min}}$
p_3	$1.3 \times 10^{-4} \frac{\text{L}}{\text{mu} \cdot \text{min}}$
n	$2.3 \times 10^{-1} \frac{1}{\text{min}}$
BW	75 kg
V_I	$0.12BW \text{ L}$
V_G	$0.16BW \text{ L}$
M_{wG}	$180.1558 \frac{\text{g}}{\text{mol}}$
A_g	0.8
C	$18.018 \frac{\text{mmol}}{\text{L}}$
τ_D	40 min
τ_S	55 min

appropriate model parameter values for scenarios A and B are given in Table 3.

Table 3: The approximate model parameters applied in Scenarios A - B.

Notation	In Scenario A	In Scenario B
G_{Ba}^\dagger	G_B	G_B
I_{Ba}^\dagger	I_B	I_B
p_{1a}	$1.2p_1$	$0.8p_1$
p_{2a}	$0.8p_2$	$1.2p_2$
p_{3a}	$0.8p_3$	$1.2p_3$
n_a	$0.8n$	$1.2n$
BW_a^\dagger	BW	BW
V_{Ia}	$0.12BW_a$	$0.12BW_a$
V_{Ga}	$0.16BW_a$	$0.16BW_a$
M_{wGa}^\dagger	M_{wG}	M_{wG}
A_{ga}^\dagger	A_g	A_g
C_a^\dagger	C	C
τ_{Da}	$0.8\tau_D$	$1.2\tau_D$
τ_{Sa}	$1.2\tau_S$	$0.8\tau_S$

[†]It was assumed that these parameters were exactly known.

The simulation results are presented in Figs. 7 - 22. Each of these figures reveals acceptable numerical values

and no any cut was necessary in the models to exclude the physically not interpretable $u < 0$ insulin ingress rates, and the negative concentration values.

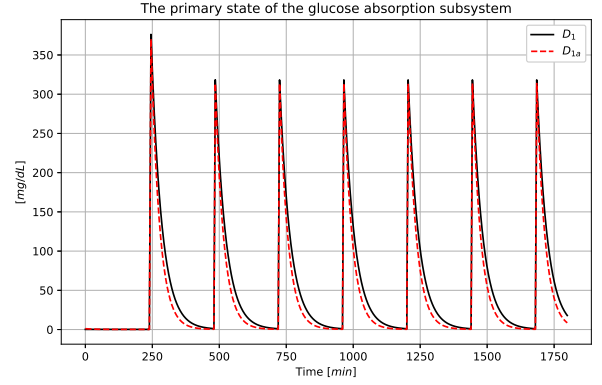


Fig. 7: The primary state variable of the glucose absorption subsystem in "Scenario A"

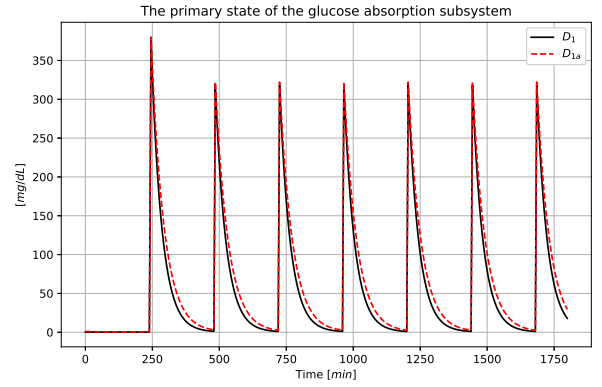


Fig. 8: The primary state variable of the glucose absorption subsystem in "Scenario B"

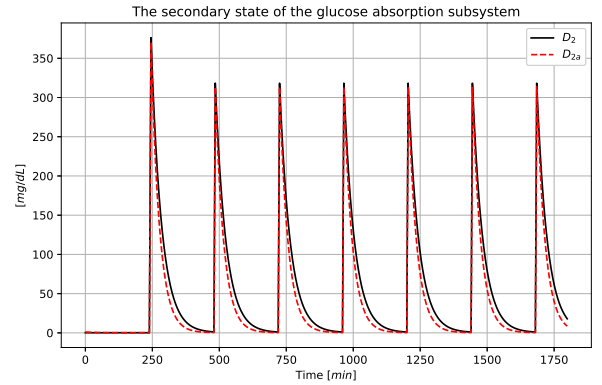


Fig. 9: The secondary state variable of the glucose absorption subsystem in "Scenario A"

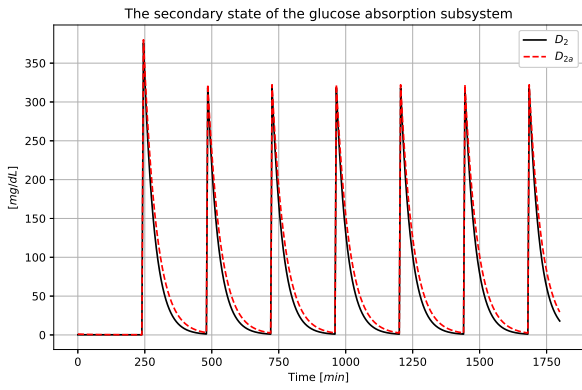


Fig. 10: The secondary state variable of the glucose absorption subsystem in "Scenario B"

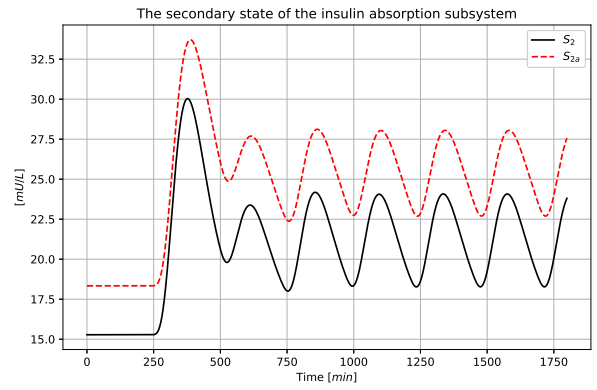


Fig. 13: The secondary state variable of the insulin absorption subsystem in "Scenario A"

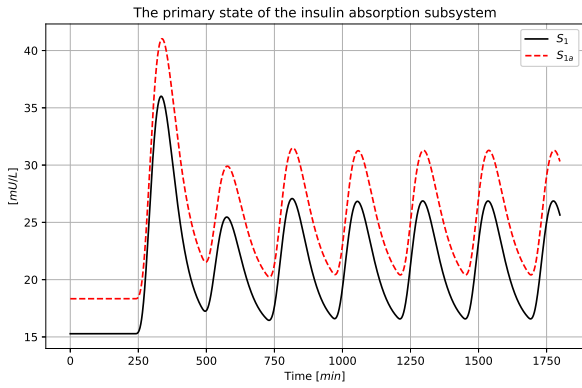


Fig. 11: The primary state variable of the insulin absorption subsystem in "Scenario A"

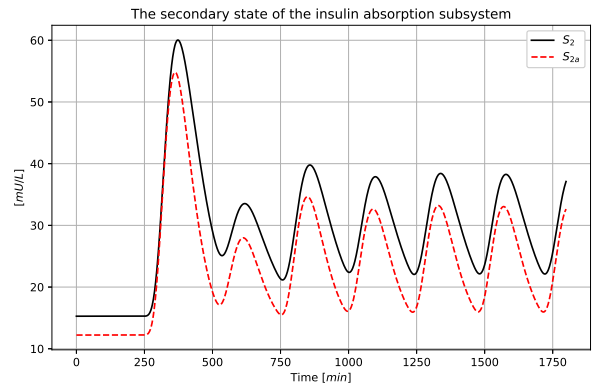


Fig. 14: The secondary state variable of the insulin absorption subsystem in "Scenario B"

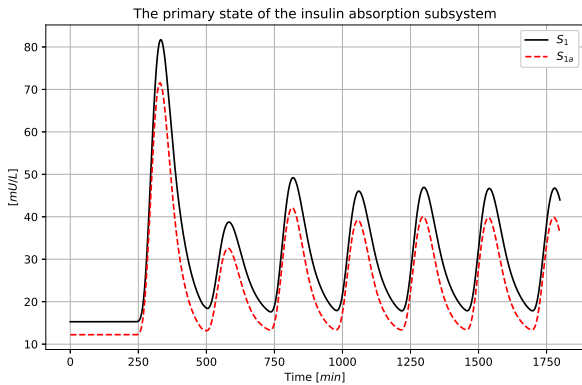


Fig. 12: The primary state variable of the insulin absorption subsystem in "Scenario B"

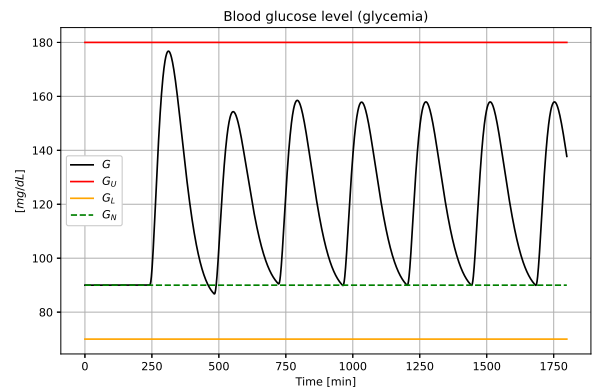


Fig. 15: The blood glucose level in "Scenario A" (G_U and G_L denote the upper and lower bounds of the risk-free region)

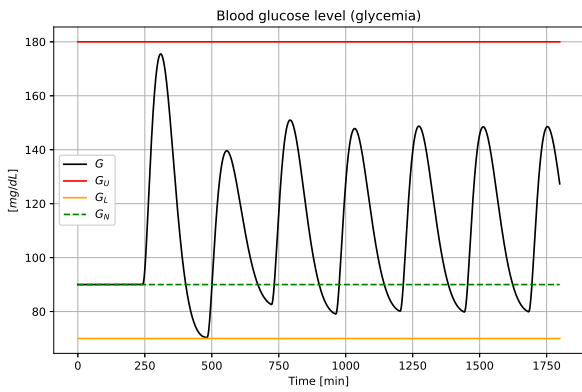


Fig. 16: The blood glucose level in "Scenario B" (G_U and G_L denote the upper and lower bounds of the risk-free region)

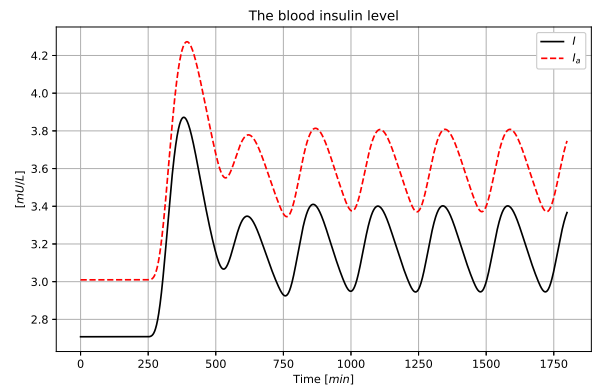


Fig. 19: The blood insulin level in "Scenario A"

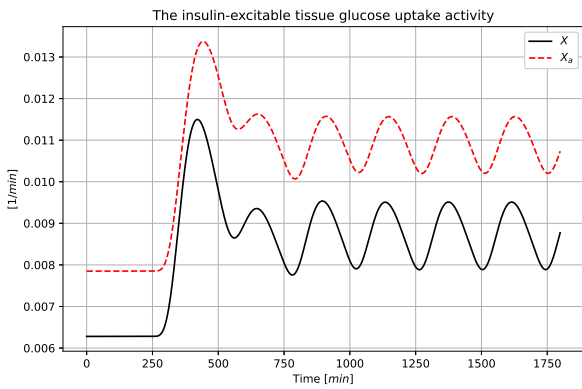


Fig. 17: The insulin-excitable tissue glucose uptake activity in "Scenario A"

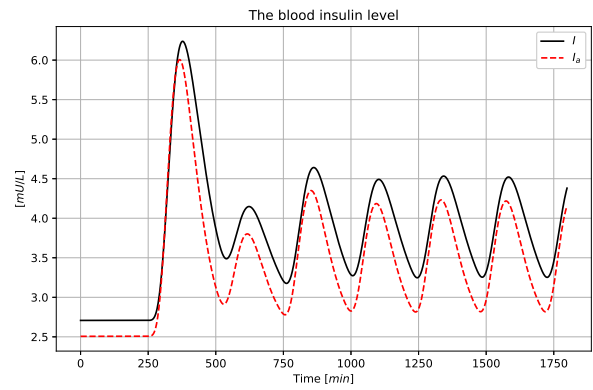


Fig. 20: The blood insulin level in "Scenario B"

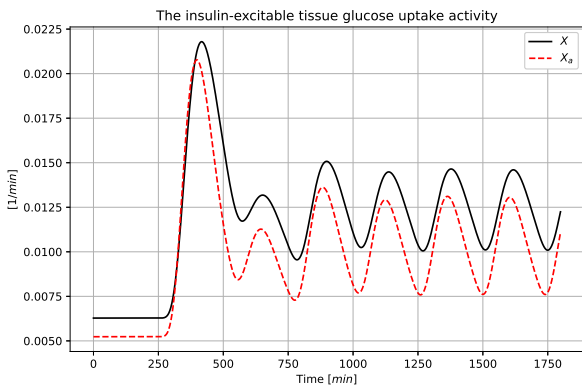


Fig. 18: The insulin-excitable tissue glucose uptake activity in "Scenario B"

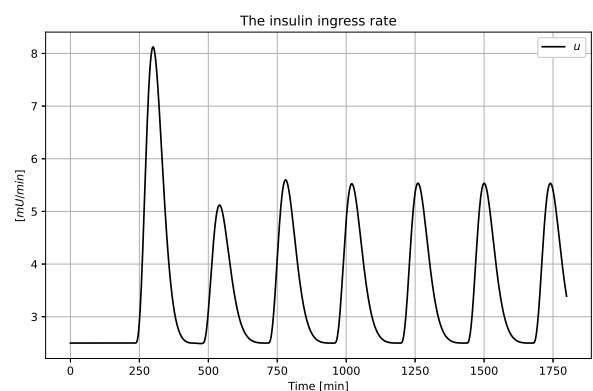


Fig. 21: The insulin ingress rate (the control signal) in "Scenario A"

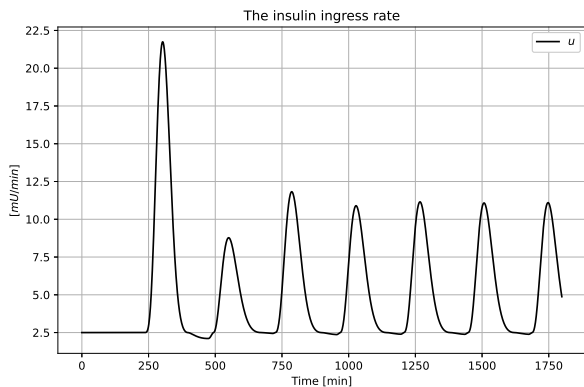


Fig. 22: The insulin ingress rate (the control signal) in "Scenario B"

It can be stated that the results obtained seem to be acceptable and it is definitely reasonable to make further calculations for various models.

V. FURTHER RESEARCH PLANS

In this paper, by the use of a simple second order mechanical paradigm it was shown that the RHC controllers can be numerically so realized that the use of Lagrange's Reduced Gradient Method can be avoided and instead it a simpler gradient descent algorithm can be run. Application examples were given for the control of a second order mechanical system and a first order model of the human glucose-insulin system to be used for treating patients suffering from type 1 diabetes mellitus. The suggested approach has the main advantage that it allows the use of arbitrarily structured cost functions. In resulting successful optimization the special form of the cost function in (16) distinguishing between the "small" and the "significant" tracking error components played an important role. It was shown that though in the lack of an appropriate sensor the not observable, "hidden" state variables were estimated only on the basis of model calculations, their estimation error at least can be kept at bay and the so estimated values can be utilized in the receding horizon control. In the future we wish to apply this method for modeling control-based T1DM treatment in which we plan to use more sophisticated human glucose-insulin system models than that used in these preliminary investigations. We also wish to use more complex COH intake disturbance systems, too. It has to be emphasized that in the present considerations the selection of the "initial state" was important, too. In the later investigations more complex initial states will be considered, too.

Contribution of individual authors to the creation of a scientific article (ghostwriting policy)

Hemza Redjimi developed the Julia codes used for the simulation, carried out the computations, and wrote a draft version of the paper.

József K. Tar checked the results and finalized the paper.

We acknowledge the support of this work by the Doctoral School of Applied Informatics and Applied Mathematics of Óbuda University.

REFERENCES

- [1] G. Pacini and R.N. Bergman. MINMOD: a computer program to calculate insulin sensitivity and pancreatic responsiveness from the frequently sampled intravenous glucose tolerance test. *Computer Methods and Programs in Biomedicine*, 23:113–122, 1986.
- [2] J.T. Sørensen. *A Physiologic Model of Glucose Metabolism in Man and Its use to Design and Assess Improved Insulin Therapies for Diabetes*. Massachusetts Institute of Technology, 1985.
- [3] C. Cobelli and G. Pacini. Insulin secretion and hepatic extraction in humans by minimal modeling of C-peptide and insulin kinetics. *Diabetes*, 37:223–231, 1988.
- [4] R.N. Bergman. Lilly lecture 1989. Toward physiological understanding of glucose tolerance. Minimal-model approach. *Diabetes*, 38:1512–1527, 1989.
- [5] T.C. Ni, M. Ader, and E.N. Bergman. Reassessment of glucose effectiveness and insulin sensitivity from minimal model analysis: a theoretical evaluation of the single-compartment glucose distribution assumption. *Diabetes*, 46:1813–1821, 1997.
- [6] R. Hovorka, F. Shojaee-Moradie, P.V. Carroll, L.J. Chassin, I.J. Gowrie, N.C. Jackson, R.S. Tudor, A.M. Umpleby, and R.H. Jones. Partitioning glucose distribution/transport, disposal, and endogenous production during ivgtt. *Am J Physiol Endocrinol Metab*, 282:E992–E1007, 2002.
- [7] C. Dalla Man, R.A. Rizza, and C. Cobelli. Mixed meal simulation model of glucose-insulin system. *Proc. of the 28th IEEE EMBC Conference, New York, USA*, pages 307–310, 2006.
- [8] C. Dalla Man, R.A. Rizza, and C. Cobelli. Meal simulation model of glucose-insulin system. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 54(10):1740–1749, 2007.
- [9] E. Friis-Jensen. *Modeling and Simulation of Glucose-Insulin Metabolism*. PhD Thesis, Technical University of Denmark, Kongens Lyngby, Denmark, 2007.
- [10] C. Dalla Man, F. Micheletto, D. Lv, M. Breton, B.P. Kovatchev, and C. Cobelli. The UVA/PADOVA Type 1 diabetes simulator: New features. *J Diab Scien Techn*, 8(8):26–34, 2014.
- [11] R. Hovorka, V. Canonico, L.J. Chassin, U. Haueter, M. Massi-Benedetti, M. Orsini-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. *Physiol Meas*, 25(4):905–920, 2004.
- [12] 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 Diab Sci Techn*, 1:804–812, 2007.
- [13] F. Chee and T. Fernando. *Closed-Loop Control of Blood Glucose*. Springer, 2007.

- [14] L. Magni, D.M. Raimondo, C. Dalla Man, M. Breton, S. Patek, G. De Nicolao, C. Cobelli, and B.P. Kovatchev. Evaluating the efficacy of closed-loop glucose regulation via control-variability grid analysis. *J Diab Scien Techn*, 2(4):630–635, 2008.
- [15] L. Kovács, B. Benyó, J. Bokor, and Z. Benyó. Induced L2-norm minimization of glucose-insulin system for type I diabetic patients. *Comp Meth Prog Biomed*, 102:105–118, 2011.
- [16] P. Herrero, P. Georgiou, N. Oliver, D.G. Johnston, and C. Toumazou. A bio-inspired glucose controller based on pancreatic β -cell physiology. *J Diab Sci Technol*, 6(3):606–616, 2012.
- [17] L. Kovács, P. Szalay, Z. Almassy, and L. Barkai. Applicability results of a nonlinear model-based robust blood glucose control algorithm. *Journal of Diabetes Science and Technology*, 7(3):708–716, 2013.
- [18] P. Szalay, Gy. Eigner, and L. Kovács. Linear matrix inequality-based robust controller design for type-1 diabetes model. *Proc. of the 19th World Congress of The International Federation of Automatic Control, Cape Town, South-Africa*, pages 9247–9252, 2014.
- [19] H. Khan, J.K. Tar, I. Rudas, L. Kovács, and Gy. Eigner. Receding Horizon Control of Type 1 Diabetes Mellitus by using Nonlinear Programming. *Complexity*, <https://doi.org/10.1155/2018/4670159>, 2018.
- [20] R. Hovorka, V. Canonico, L.J. Chassin, U. Haueter, M. Massi-Benedetti, M. Orsini-Federici, T.R. Pieber, H.C. Schaller, L. Schaupp, T. Vering, and Wilinska M.E. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol Meas*, 25(4):905 – 920, 2004.
- [21] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control*. Springer, 2011.
- [22] J. Richalet, A. Rault, J.L. Testud, and J. Papon. Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14(5):413–428, 1978.
- [23] J.L. Lagrange, J.P.M. Binet, and J.G. Garnier. *Mécanique analytique (Eds. J.P.M. Binet and J.G. Garnier)*. Ve Courcier, Paris, 1811.
- [24] Tjalling J. Ypma. Historical development of the Newton-Raphson method. *SIAM Review*, 37(4):531–551, 1995.
- [25] C.T. Kelley. *Solving Nonlinear Equations with Newton’s Method, no 1 in Fundamentals of Algorithms*. SIAM, 2003.
- [26] J. Houston. Economic optimisation using Excel’s Solver: A macroeconomic application and some comments regarding its limitations. *Computers in Higher Education Economics Review*, 11(1):2–5, 1997.
- [27] E. Nævdal. Solving continuous-time optimal-control problem with a spreadsheet. *Journal of Economic Education*, 34(2):99–122, 2003.
- [28] H.B. Callen. *Thermodynamics and an Introduction to Thermostatistics (Second Edition)*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, 1960.
- [29] H. Karabulut. Physical meaning of Lagrange multipliers. *European Journal of Physics (physics.ed-ph); General Physics (physics.gen-ph)*, 27:709–718, 2007.
- [30] V.I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer - Verlag, 1989.
- [31] R.E. Kalman. Contribution to the theory of optimal control. *Boletín Sociedad Matemática Mexicana*, 5(1):102–119, 1960.
- [32] J. Riccati. Animadversiones in aequationes differentiales secundi gradus (observations regarding differential equations of the second order). *Actorum Eruditorum, quae Lipsiae publicantur, Supplementa*, 8:66–73, 1724.
- [33] A.J. Laub. *A Schur Method for Solving Algebraic Riccati Equations (LIDS-P 859 Research Report)*. MIT Libraries, Document Services, After 1979.
- [34] Zs. Horváth and A. Edelmayer. Solving the modified filter algebraic Riccati equation for H-infinity fault detection filtering. *Acta Universitatis Sapientiae, Electrical and Mechanical Engineering*, 9:57–77, 2017.
- [35] W.M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM Journal on Control and Optimization*, 6(1):681–697, 1968.
- [36] Hazem Issa and J.K. Tar. Speeding up the Reduced Gradient Method for constrained optimization. *Accepted for publication IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMII 2021), January 21-23, Herl’any, Slovakia*, pages 1–6, 2021.
- [37] J.P. Gram. Über die Entwicklung reeler Funktionen in Reihen mittelst der Methode der kleinsten Quadrate. *Journal für die reine und angewandte Mathematik*, 94:71–73, 1883.
- [38] E. Schmidt. Zur Theorie der linearen und nichtlinearen Integralgleichungen. I. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener. *Mathematische Annalen*, 63:442, 1907.
- [39] P.-S. Laplace. *Théorie Analytique des Probabilités, in: Oeuvres complètes de Laplace / publiés sous les auspices de l’Académie des sciences, par MM. les secrétaires perpétuel*. Gauthier-Villars, Paris, 1886.