Optimal Insulin Administration for People with Type 1 Diabetes

Dimitri Boiroux ∗ Daniel A. Finan ∗ John Bagterp Jørgensen ∗ Niels Kjølstad Poulsen ∗ Henrik Madsen ∗

∗ Department of Informatics and Mathematical Modeling, Technical University of Denmark, DK-2800 Kgs Lyngby, Denmark (e-mail: {dibo, dafi, jbj, nkp, hm}@imm.dtu.dk)

Abstract: In this paper we apply receding horizon constrained optimal control to the computation of insulin administration for people with type 1 diabetes. The study is based on the Hovorka model, which describes a virtual subject with type 1 diabetes. First of all, we compute the optimal insulin administration for the linearized system using quadratic programming (QP) for optimization. The optimization problem is a discrete-time problem with soft state constraints and hard input constraints. The computed insulin administration is applied to the nonlinear model, which represents the virtual patient. Then, a nonlinear discrete-time Bolza problem is stated and solved using sequential quadratic programming (SQP) for optimization and an explicit Dormand-Prince Runge-Kutta method (DOPRI54) for numerical integration and sensitivity computation. Finally, the effects of faster acting insulin on the postprandial (i.e., post-meal) blood glucose peak are discussed.

Keywords: Type 1 diabetes, model predictive control, physiological modeling, simulation

1. INTRODUCTION

The World Health Organization (2008) estimates that 180 million people worldwide have diabetes. This number is predicted to double by 2030. In the USA, the budget for diabetes alone represents 10% of the health care budget, or more than 130 billion dollars (132 billion dollars in 2002).

In people without diabetes, the blood glucose is controlled tightly around 90 mg/dL (~5 mmol/L). Type 1 diabetes is a chronic disease characterized by an insufficient (effectively nonexistent) endogenous production of insulin, which leads to poor regulation of glucose concentrations in the blood. In particular, the deficiency of insulin causes sustained high glucose levels (hyperglycemia) that result in serious long-term health problems like eye, nerve, and kidney disease. On the other hand, too much insulin can result in very low glucose levels (hypoglycemia) which can pose immediate health risks. Exogenous insulin, then, must be injected to regulate the blood glucose concentration as tightly as possible.

Usually, insulin treatment consists of the administration of rapid acting insulin through boluses (i.e., discrete insulin administration) at the time of meals. The size of the bolus is based on a measurement of the current blood glucose at mealtime and the (estimated) size of the meal. However, having measurements only at mealtime does not provide enough information about blood glucose. Hypoglycemic and hyperglycemic events can go unobserved due to the infrequent blood glucose measurements. In addition, such a measurement process does not give any information about the dynamic trend of the blood glucose. Consequently, people with diabetes often tolerate hyperglycemia in order to avoid hypoglycemia and its immediate effects.

Continuous glucose monitors (CGM) can help to provide a better control of blood glucose. They measure the glucose concentration in the subcutaneous depot. Insulin pumps that continuously inject fast acting insulin have also been developed. Combining a CGM with an insulin pump can enable automatic insulin administration for people with type 1 diabetes. Such a medical device is called an artificial pancreas and is illustrated in Fig. 1. Several research groups work on aspects of control algorithms integrating the CGM and the insulin pump to automatically adjust insulin administration for people with type 1 diabetes (see e.g. Klonoff et al. (2009)).

In this paper we describe the Hovorka model and use this description to point to the factors limiting ideal glucose
control by insulin administration. One factor limiting the performance is the relative long absorption time of insulin. Using open-loop NMPC we describe quantitatively the maximal postprandial glucose in relation to the insulin absorption rate.

The paper is structured as following. Section 2 presents the model developed by Hovorka et al. (2004). Section 3 states an optimal control problem in the linear case. Section 4 presents the nonlinear optimal control problem and discusses the benefits of having faster-acting insulin. Conclusions are provided in Section 5.

2. MODEL

For the study of insulin administration and its effect on glucose concentrations we use a model developed by Hovorka et al. (2002, 2004). The model consists of a submodel describing food absorption, a submodel describing subcutaneous-to-intravenous absorption of insulin, a simple lumped model describing the glucose dynamics, and simple lumped models describing insulin dynamics and action mechanisms. In the following we describe these models.

2.1 Food Absorption

Food absorption models have been considered by a number of authors (Elashoff et al., 1982; Lehmann and Deutsch, 1992; Dalla Man et al., 2006; Goetze et al., 2007) and it has been observed that people with diabetes has abnormally slow gastric emptying (Horowitz et al., 2002).

In this paper, we consider a two-compartment model describing carbohydrate (CHO) absorption and conversion to glucose. The model describes the effect of orally ingested carbohydrates on the rate of appearance of glucose in the blood stream. The model is

\[
\begin{align*}
\frac{dD_1}{dt}(t) &= A_G D(t) - \frac{1}{\tau_D} D_1(t) \\
\frac{dD_2}{dt}(t) &= \frac{1}{\tau_D} D_1(t) - \frac{1}{\tau_D} D_2(t)
\end{align*}
\]

in which \( D(t) \) [mmol/min] is the amount of oral carbohydrate intake at any time expressed as glucose equivalents, \( A_G \) is a factor describing the utilization of carbohydrates to glucose, \( \tau_D \) [min] is the time constant, \( D_1(t) \) [mmol] and \( D_2(t) \) [mmol] are the states describing the amount of glucose in the two compartments. The rate of appearance of absorption of glucose in the blood stream is described by

\[
U_G(t) = \frac{1}{\tau_D} D_2(t)
\]

\( U_G(t) \) [mmol/min] is the glucose absorption rate. The carbohydrate input rate, \( D(t) \) [mmol/min], may be related to the carbohydrate input rate, \( d(t) \) [g/min], by

\[
D(t) = \frac{1000}{M_wG} d(t)
\]

in which \( M_wG \) [g/mol] is the molecular weight of glucose.

2.2 Insulin Absorption

Insulin is administered subcutaneously. A number of models to describe the absorption rate of subcutaneously injected short acting insulin in the blood stream are available (Wilinska et al., 2005).

In this paper we consider a two compartment model describing the absorption rate of subcutaneously administered short acting insulin. The model is

\[
\begin{align*}
\frac{dS_1}{dt}(t) &= u(t) - \frac{1}{\tau_S} S_1(t) \tag{4a} \\
\frac{dS_2}{dt}(t) &= \frac{1}{\tau_S} S_1(t) - \frac{1}{\tau_S} S_2(t) \tag{4b}
\end{align*}
\]

in which \( u(t) \) [mU/min] is the amount of insulin injected, \( \tau_S \) [min] is the time constant, \( S_1(t) \) [mU] and \( S_2(t) \) [mU] are the amounts of insulin in the two compartments. The absorption rate of insulin in the blood stream is

\[
U_I(t) = \frac{1}{\tau_S} S_2(t)
\]

in which \( U_I(t) \) [mU/min] is the absorption rate.

2.3 Glucose Subsystem

The blood glucose dynamics are modeled with two compartments. The two state variables are \( Q_1(t) \) [mmol] and \( Q_2(t) \) [mmol]. \( Q_1(t) \) represents glucose in the main blood stream, while \( Q_2(t) \) represents glucose in peripheral tissue such as muscles.

The model describing evolution of glucose in the main blood stream

\[
\begin{align*}
\frac{dQ_1}{dt}(t) &= U_G(t) - F_{01,c}(t) - F_R(t) - x_1(t) Q_1(t) + k_{12} Q_2(t) + EGP_0(1 - x_3(t)) \tag{6}
\end{align*}
\]

includes absorption from the gut, \( U_G(t) \) [mmol/min], consumption of glucose by the central nervous system, \( F_{01,c} \) [mmol/min], the renal excretion of glucose in the kidneys, \( F_R(t) \) [mmol/min], the insulin dependent uptake of glucose in muscles, \( x_1(t) Q_1(t) \) [mmol/min], transfer of glucose from peripheral tissue such as muscle to the blood, \( k_{12} Q_2(t) \), and endogenous release of glucose by the liver, \( EGP_0(1 - x_3(t)) \). The uptake of glucose in muscles depends on insulin. \( x_1(t) \) is a state representing insulin in muscle tissue. Release of glucose from the liver is also controlled by insulin. High concentrations of insulin suppress glucose release. \( x_3(t) \) is used to model insulin in the liver.
Glucose in peripheral tissue such as muscle is modeled by the differential equation

$$\frac{dQ_2}{dt}(t) = x_1(t)Q_1(t) - (k_{12} + x_2(t))Q_2(t)$$  \hspace{1cm} (7)

in which $x_1(t)Q_1(t)$ [mmol/min] is the transport of glucose from the main blood stream to the muscles, $k_{12}Q_2(t)$ [mmol/min], is transport of peripheral glucose to the main blood stream, and $x_2(t)Q_2(t)$ [mmol/min] is the insulin dependent disposal of glucose in the muscle cells. It depends on insulin modeled by $x_2(t)$.

The glucose concentration is

$$y(t) = G(t) = \frac{Q_1(t)}{V_G}$$ \hspace{1cm} (8)

$y(t) = G(t)$ is the glucose concentration [mmol/L] and $V_G$ is the glucose distribution volume. It depends on body weight, $BW$ [kg], of the individual.

The consumption of glucose by the central nervous systems is modeled as

$$F_{01,c}(t) = \begin{cases} F_{01} & G(t) \geq 4.5 \text{ mmol/L} \\ F_{01}G(t)/4.5 & \text{otherwise} \end{cases}$$  \hspace{1cm} (9)

At low glucose concentrations the consumption, $F_{01,c}$ [mmol/min], is proportional to the glucose concentration, $G(t)$, while it is constant when the glucose concentration is not low.

The excretion rate of glucose in the kidneys is zero unless the glucose concentration is high ($G(t) \geq 9$ mmol/L). In this case it is affine in the glucose concentration. Consequently, the glucose excretion rate, $F_R$ [mmol/min], is modeled as

$$F_R(t) = \begin{cases} 0.003(G(t) - 9)V_G & G(t) \geq 9 \text{ mmol/L} \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (10)

### 2.4 Insulin Subsystem

Then the plasma insulin concentration, $I(t)$ [mU/L], evolves according to

$$\frac{dI}{dt}(t) = \frac{U_I(t)}{V_I} - k_eI(t)$$  \hspace{1cm} (11)

The insulin action is governed by influence on transport and distribution $x_1(t)$, utilization and phosphorylation of glucose in adipose tissue $x_2(t)$, and endogenous production in the liver $x_3(t)$. These quantities are described by the differential equations

$$\frac{dx_1}{dt}(t) = -k_{a1}x_1(t) + k_{b1}I(t)$$  \hspace{1cm} (12a)

$$\frac{dx_2}{dt}(t) = -k_{a2}x_2(t) + k_{b2}I(t)$$  \hspace{1cm} (12b)

$$\frac{dx_3}{dt}(t) = -k_{a3}x_3(t) + k_{b3}I(t)$$  \hspace{1cm} (12c)

### 2.5 Parameters

The parameters in the Hovorka model (1)-(12) are listed in Table 1. The parameters $k_{b,i}$ are related to the insulin sensitivities, $S_{1,i}$, by

$$k_{b,i} = S_{1,i}k_{a,i} \hspace{1cm} i = 1, 2, 3$$  \hspace{1cm} (13)

### Table 1. Parameters in the Hovorka Model

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transfer rate</td>
<td>$k_{12}$</td>
<td>0.066 1/min</td>
</tr>
<tr>
<td>Deactivation rate</td>
<td>$k_{a1}$</td>
<td>0.006 1/min</td>
</tr>
<tr>
<td>Deactivation rate</td>
<td>$k_{a2}$</td>
<td>0.06 1/min</td>
</tr>
<tr>
<td>Deactivation rate</td>
<td>$k_{a3}$</td>
<td>0.03 1/min</td>
</tr>
<tr>
<td>Insulin elimination rate</td>
<td>$k_e$</td>
<td>0.138 1/min</td>
</tr>
<tr>
<td>CHO absorption constant</td>
<td>$\tau_D$</td>
<td>40 min</td>
</tr>
<tr>
<td>Insulin absorption constant</td>
<td>$\tau_I$</td>
<td>55 min</td>
</tr>
<tr>
<td>CHO utilization</td>
<td>$A_C$</td>
<td>0.8</td>
</tr>
<tr>
<td>Transport insulin sensitivity</td>
<td>$S_{1,1}$</td>
<td>$51.2 \cdot 10^{-4}$ L/mU</td>
</tr>
<tr>
<td>Disposal insulin sensitivity</td>
<td>$S_{1,2}$</td>
<td>$8.2 \cdot 10^{-4}$ L/mU</td>
</tr>
<tr>
<td>EGP insulin sensitivity</td>
<td>$S_{1,3}$</td>
<td>$520 \cdot 10^{-4}$ L/mU</td>
</tr>
<tr>
<td>Glucose distribution volume</td>
<td>$V_G$</td>
<td>0.12 L/kg</td>
</tr>
<tr>
<td>Glucose distribution volume</td>
<td>$V_{GW}$</td>
<td>0.16 L/kg</td>
</tr>
<tr>
<td>Liver glucose production</td>
<td>$EGP_0$</td>
<td>$0.0161 \text{ mmol/kg}$</td>
</tr>
<tr>
<td>CNS glucose consumption</td>
<td>$U_{CNS}$</td>
<td>$0.0097 \text{ mmol/kg}$</td>
</tr>
<tr>
<td>CHO absorption constant</td>
<td>$\tau_D$</td>
<td>40 min</td>
</tr>
<tr>
<td>Insulin elimination rate</td>
<td>$k_e$</td>
<td>0.138 1/min</td>
</tr>
<tr>
<td>CHO utilization</td>
<td>$A_C$</td>
<td>0.8</td>
</tr>
<tr>
<td>Transport insulin sensitivity</td>
<td>$S_{1,1}$</td>
<td>$51.2 \cdot 10^{-4}$ L/mU</td>
</tr>
<tr>
<td>Disposal insulin sensitivity</td>
<td>$S_{1,2}$</td>
<td>$8.2 \cdot 10^{-4}$ L/mU</td>
</tr>
<tr>
<td>EGP insulin sensitivity</td>
<td>$S_{1,3}$</td>
<td>$520 \cdot 10^{-4}$ L/mU</td>
</tr>
<tr>
<td>Glucose distribution volume</td>
<td>$V_G$</td>
<td>0.12 L/kg</td>
</tr>
<tr>
<td>Glucose distribution volume</td>
<td>$V_{GW}$</td>
<td>0.16 L/kg</td>
</tr>
<tr>
<td>Liver glucose production</td>
<td>$EGP_0$</td>
<td>$0.0161 \text{ mmol/kg}$</td>
</tr>
<tr>
<td>CNS glucose consumption</td>
<td>$U_{CNS}$</td>
<td>$0.0097 \text{ mmol/kg}$</td>
</tr>
</tbody>
</table>

### 3. LINEAR MODEL PREDICTIVE CONTROL

In this section, we formulate and discuss the linearized optimal control problem. Let $x(t) \in \mathbb{R}^{n_x}$ be the state vector, $u(t) \in \mathbb{R}^{n_u}$ be the manipulated inputs, and $d(t) \in \mathbb{R}^{n_d}$ be known disturbances.

A zero-order hold parametrization for the manipulated variables function $u$ and the disturbance function $d$ is used. We divide the time interval, $[t_0, t_f]$, into $N$ equidistant intervals, each of length $T_a$. We denote $N = \{0, 1, ..., N - 1\}$ and $t_k = t_0 + kT_a$ for $k \in N$. The zero-order hold restrictions on $d$ and $u$ imply

$$u(t) = u_k \hspace{1cm} t_k \leq t < t_{k+1} \hspace{1cm} k \in N \hspace{1cm} (15a)$$

$$d(t) = d_k \hspace{1cm} t_k \leq t < t_{k+1} \hspace{1cm} k \in N \hspace{1cm} (15b)$$

### 3.1 Hard Output Constraints

Using the zero-order hold parametrization (15), the linear discrete-time optimal control problem may be expressed as
\[
\begin{align*}
\min_{\{w_k\}_{k=0}^{N-1}} \phi &= \phi \left( \{u_k, y_{k+1}, r_{k+1}\}_{k=0}^{N-1} \right) \\
\text{s.t.} \quad &x_{k+1} = Ax_k + Bu_k + Ed_k \\
&y_k = Cx_k \\
&u_{\min} \leq u_k \leq u_{\max} \\
&\Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max} \\
&y_{\min} \leq y_k \leq y_{\max} \\
&v_k \geq 0 \\
&w_k \geq 0
\end{align*}
\]

in which \(x_k \in \mathbb{R}^{n_x}\) is the state vector at time \(t_k\), and \(y_k\) is the measured output at time \(t_k\). The manipulated inputs \(u_k\) and the difference \(\Delta u_k = u_{k+1} - u_k\) must lie in the interval \([u_{\min}, u_{\max}]\) and \([\Delta u_{\min}, \Delta u_{\max}]\) respectively.

The objective function \(\phi\) is in the form
\[
\phi = \frac{1}{2} \sum_{k=0}^{N-1} \|y_{k+1} - r_{k+1}\|_2^2 + \lambda \|\Delta u_k\|_2^2
\]

Furthermore, the measured output \(y_k\) must lie in the interval \([y_{\min}, y_{\max}]\). In the insulin administration problem, it means that the blood glucose at sample times must be kept in the normoglycemic range (60-140 mg/dL or 3.3-7.8 mmol/L).

However, infeasibility of (16) might arise due to the hard constraints on the measured output \(y_k\). Consequently, it is preferable to replace the hard constraints (16f) with soft output constraints.

### 3.2 Soft Output Constraints

The hard constraints (16f) are replaced by soft constraints using the slacks variables \(v_k\) and \(w_k\). The new objective function is
\[
\phi = \frac{1}{2} \sum_{k=0}^{N-1} \|y_{k+1} - r_{k+1}\|_2^2 + \lambda \|\Delta u_k\|_2^2
\]

\[
\kappa_1 \|v_k\|_2^2 + \kappa_2 \|w_k\|_2^2
\]

The two new terms \(\kappa_1 \|v_k\|_2^2 + \kappa_2 \|w_k\|_2^2\) correspond to penalty costs for hyperglycemia and hypoglycemia respectively.

The linear discrete-time optimal control problem with soft constraints that has to be solved may be expressed as
\[
\min_{\{u_k, v_k, w_k\}_{k=0}^{N-1}} \phi \left( \{u_k, y_{k+1}, r_{k+1}, v_k, w_k\}_{k=0}^{N-1} \right)
\]

\[
\text{s.t.} \quad x_{k+1} = Ax_k + Bu_k + Ed_k \\
y_k = Cx_k \\
u_{\min} \leq u_k \leq u_{\max} \\
\Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max} \\
y_{\min} \leq y_k \leq y_{\max} \\
v_k \geq 0 \\
w_k \geq 0
\]

in which the hard output constraint (16f) has been replaced with penalty terms in the objective function (18) and the inequality constraints (19f - 19i).

### 3.3 Linear simulation results

We use the Hovorka et al. (2004) model linearized at the steady state corresponding to the target blood glucose concentration \(\bar{G} = 5\) mmol/L to compute the optimal insulin administration profiles. Then, we apply this profile to the Hovorka et al. (2004) model.

The objective of the insulin administration is to compensate glucose excursions caused by meals and variations in endogenous glucose production and utilization. We use a penalty function defined by (18). \(y_k\) is the blood glucose concentration, \(r_k = 5\) mmol/L is the target value for the blood glucose concentration, \(y_{\min} = 4\) mmol/L is a lower acceptable limit on the glucose concentration, and \(y_{\max} = 8\) mmol/L is an upper acceptable limit on the blood glucose concentration. The weights \(\kappa_1\) and \(\kappa_2\) are used to balance the desirability of different deviations from the target. As hypoglycemia is considered a more immediate danger than hyperglycemia, \(\kappa_1 < \kappa_2\).

The choice of the weight \(\lambda\) should not change the shape of the optimal blood glucose profile. It is used to avoid ill-conditioning of the problem. For all the simulations, we use \(\lambda = 10^{-2}\).

We use \(u_{\min} = 0\) and a large \(u_{\max}\) such that the upper bound is never active. We do the optimization in a 24 hour window, i.e. \(t_0 = 0\) min and \(t_f = 24 - 60\) min, using a sampling time of \(T_s = 5\) min. In the three scenarios considered, the simulated 70 kg subject has a meal at 6:00. The meal sizes for each scenario are 25 g CHO, 50 g CHO and 100 g CHO, respectively. We compute the optimal insulin administration using the linearized model, and simulate a virtual patient using this sequence of insulin administration on the Hovorka model (Hovorka et al. (2004)).

Fig. 3 illustrates an optimal insulin administration profile for the case where the meal size is relatively small. Having a small meal implies a small deviation to the steady state. Consequently, the linear and nonlinear solutions are quite similar.
In this paper, we described a model developed by Hovorka et al. (2004) to study the effects of insulin administration on glucose concentration for people with type 1 diabetes. Based on a linearized version of this model, we use an nonlinear model predictive control to compute the optimal insulin injection profiles for people with type 1 diabetes. The controller can deliver no anticipatory insulin prior to the meal, and the final time, \( t_f \), are specified parameters. The initial state, \( x_0 \), is a known parameter in (20). The inputs are bound-constrained and must lie in the interval \([u_{\text{min}}, u_{\text{max}}]\), and the difference \(\Delta u_k = u_{k+1} - u_k\) must lie in the interval \([\Delta u_{\text{min}}, \Delta u_{\text{max}}]\).

The objective of the insulin administration is to mitigate glucose excursions caused by meals and variations in endogenous glucose production and utilization. We use a penalty function defined as

\[
\phi = \phi \left( \{u(t), y(t), r(t)\}_{t=t_0}^{t=t_f} \right)
\]

subject to

\[
\begin{align*}
\dot{x}(t) &= f(x(t), u(t), d(t)) \\
y(t) &= g(x(t)) \\
u(t) &= u_k \\
\end{align*}
\]

for small-sized meals (25 g CHO), normal-sized meals (50 g CHO) and large-sized meals (100 g CHO) if the meal is announced only at mealtime. A faster insulin reduces the peak of glucose. For normal-sized meals, having an insulin absorption time constant at least equal to the glucose absorption time constant (i.e. \(\tau_s = 40\) minutes) avoids hypoglycemic events.

4. NONLINEAR MODEL PREDICTIVE CONTROL

In this section, we state and discuss the continuous-time nonlinear optimal control problem that we use to compute the insulin injection profiles for people with type 1 diabetes. The bound-constrained continuous-time optimal control problem is used to compute the optimal insulin administration. The mismatch between the linear and the nonlinear model becomes more evident, hypoglycemia is avoided.

Fig. 4 shows the maximum blood glucose versus the insulin time constant \(\tau_s\) for small-sized meals (25 g CHO), normal-sized meals (50 g CHO) and large-sized meals (100 g CHO) if the meal is announced only at mealtime. A faster insulin reduces the peak of glucose. For normal-sized meals, having an insulin absorption time constant at least equal to the glucose absorption time constant (i.e. \(\tau_s = 40\) minutes) avoids hypoglycemic events.

5. CONCLUSION

In this paper, we described a model developed by Hovorka et al. (2004) to study the effects of insulin administration on glucose concentration for people with type 1 diabetes. Based on a linearized version of this model, we use an...
Fig. 6. Optimal insulin administration profile obtained using NMPC.

Fig. 7. Optimal insulin administration with meal announcement at meal time.

We use our optimal control algorithm to compute insulin administration profiles for the cases with and without meal announcement in advance. The insulin profile simulated on the nonlinear model does not match the optimal insulin administration for the linearized model only if the meal is too large.

We use our optimal control algorithm to compute insulin administration profiles for the cases with and without meal announcement in advance. The results suggest that having faster acting insulin can significantly increase the control quality of blood glucose.

REFERENCES


