Identification for Control of Biomedical Systems using a very Short Experiment

Soltesz, Kristian; Mercader, Pedro

Published in:
2016 International Conference on Systems in Medicine and Biology (ICSMB)

DOI:
10.1109/ICSMB.2016.7915107

2016

Document Version:
Peer reviewed version (aka post-print)

Link to publication

Citation for published version (APA):
Soltesz, K., & Mercader, P. (2016). Identification for Control of Biomedical Systems using a very Short Experiment. In 2016 International Conference on Systems in Medicine and Biology (ICSMB) (pp. 140-143). IEEE--Institute of Electrical and Electronics Engineers Inc.. DOI: 10.1109/ICSMB.2016.7915107

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Identification for Control of Biomedical Systems using a very Short Experiment

Kristian Soltesz  
Department of Automatic Control  
Lund University  
Lund, Sweden  
Email: kristian@control.lth.se

Pedro Mercader  
Department of Computer Sciences and Systems  
University of Murcia  
Murcia, Spain  
Email: pedro.mercader@um.es

Abstract—This paper presents a combined experiment and identification procedure, well suited to obtain low-order dynamic models of a patients’ response to continuous drug administration. The experiment requires no a priori information and is of very short duration. The identification method provides both a parametric low-order model, and an estimate of the parameter error covariance. It has been demonstrated to work well with very noisy measurements, as typically encountered in drug dosing applications.

Keywords—Medical control systems, System Identification, Uncertain systems

I. INTRODUCTION

Closed-loop controlled drug delivery is becoming a reality both in anesthesia (control of hypnotic depth and analgesia) and diabetes (control of blood sugar level). There exist several prototype systems, see [5, 7] for surveys, and it is realistic to believe that these technologies will meet broad clinical acceptance within a near future. In essence these systems function according to the block diagram shown in Figure 1.

For the anesthesia case, the control signal is the administration rate of an intravenously infused drug, such as propofol, and the measurement is typically an index reflecting consciousness, derived from EEG measurements. For the diabetes case, the control signal is the insulin infusion rate, while the blood glucose level is being measured.

The technical aspect that foremost limits the development of closed-loop drug delivery systems, is the availability of reliable patient models, dynamically relating drug infusion to clinical effect. For identification of such models to be successful, the patient(s) to be modeled need to be exposed to changes in the input (drug infusion rate), while the output (clinical effect) is measured. Clinical practice and ethics limit the amount of admissible excitation in such experiments, both with respect to input signal activity and duration.

A quick review of typical model structures, and a motivation for the use of low-order approximations, is given in Section II. In Section III a short duration experiment, with limited activity in the control signal, is proposed. The use of the experiment outcome to identify low-order models, including uncertainty descriptions, is the topic of Section IV. The combination of experiment and parameter estimation method is demonstrated through a realistic example in Section V. Results are briefly discussed in Section VI.

II. PATIENT MODELS

For control purposes, the static output nonlinearity of the PD model is typically handled either by linearization close to the intended operating point [12], or an inverting gain schedule [8]. The fact that the LTI part of the PKPD model lacks oscillatory modes (due to the compartment structure) allows
A method based on closing a negative feedback loop over
the plant to be modeled, in series with a relay nonlinearity,
\( u = -\text{sgn}(y)u_{\text{on}} \), as shown in Figure 3a, was first presented
in [2]. The inverse describing function of the relay intersects
the plant in the frequency domain interpretation.

The plant input \( u \) and output \( y \), sampled at period \( h \), are
used to obtain parameter estimates \( \theta = [b \ a \ L]^T \) corresponding
to the assumed FOTD model structure
\[
\hat{P}(s) = \frac{b}{s + a} e^{-sL}.
\]

This is done by a version of the output error method used
in [11], presented below for the more general model structure
\[
\hat{P}(s) = \frac{b_1 s^{m-1} + b_2 s^{m-2} + \cdots + b_m}{s^k} e^{-sL},
\]
parameterized in \( \theta = [b^T \ a^T \ L]^T \), where \( b = [b_1 \ \ldots \ b_m]^T \)
and \( a = [a_1 \ \ldots \ a_n]^T \).

Continuous time models are used to limit the number of
elements of \( \theta \), in presence of the delay \( L \). The objective is to minimize (half the squared) \( L_2 \)-norm of the output error
e = \( \hat{y} - \tilde{y} \):
\[
J(\theta) = \frac{1}{2} \int_0^\infty e^2(t)dt,
\]
where \( \tilde{y} \) is the resulting output when \( \hat{P} \) (parameterized in \( \theta \)) is
driven by \( u \). The optimization problem is approached with a
trust-region method [4]. To improve convergence, the method
is provided with the parameter sensitivity gradient \( \nabla J \) and
Hessian \( \Delta J \). The gradient w.r.t. \( \theta \) is given by
\[
\nabla J = \int_0^\infty e(t)\nabla \hat{y}(t)dt,
\]
and the Hessian is
\[
\Delta J = \int_0^\infty \nabla \hat{y} \nabla \hat{y}^T + e(t)\Delta \hat{y}dt.
\]

The first term of the integrand in (5) is quadratic (\( \geq 0 \)),
while the integral of the second term is small (\( \approx 0 \)), under
the realistic assumption that the output error is uncorrelated
with its second derivative (\( Ee\Delta \hat{y} = 0 \)). It is therefore fair
to approximate the Hessian by only the first term (although it is straightforward to extend the method outlined below, to include also the second term). In order to account for the $k$ explicit integrators in (2), $k$ zeros are appended to $a$, forming $	ilde{a} = [a^T 0_{1 \times k}]^T$, while $b$ is padded by leading zeros to make the same length: $\tilde{b} = [0_{1 \times n-m+k} b^T]^T$. Using the results from [11] it is then possible to construct a continuous time LTI state space system, with output $[\tilde{y} \tilde{\gamma} \tilde{\omega}]^T$, when driven by $u$. From $y$, $\tilde{y}$, and $\tilde{\gamma}$, it is thereafter straightforward to compute $J$, $\nabla J$ and (the mentioned approximation of) $\Delta J$. The results of these computations are supplied in each iteration of a trust-region optimization algorithm (invoked from the Matlab fmincon command) to find the optimum $\tilde{J}$ and corresponding (expected) parameter vector $\tilde{\theta}$.

B. Parametric uncertainty

In addition to the expectation $\tilde{\theta}$, the optimization provides the asymptotic covariance matrix

$$R_{\theta} = \mathbb{E}((\theta - \tilde{\theta})(\theta - \tilde{\theta})^T) = \frac{2}{N} J(\Delta J)^{-1},$$

(6)

where $N$ is the number of samples. The standard deviations of parameter estimates decreases proportional to $\sqrt{N}$, meaning that one cannot expect significantly improved estimation precision, merely by small increases in experiment duration.

C. Notes on convergence

In previous work [3, 11], relay experiments were used to obtain reasonable initial parameters for identification schemes similar to the one presented above in Section IV. As mentioned, this requires long experiment duration for the limit cycle oscillation to converge, while the use of extremum values in the data makes the procedure sensitive to noise. By evaluation on a set of $10^4$ fourth order compartment models with random parameters, it turned out that initialization of the trust-region algorithm with the parameter vector $\theta = 0_{3 \times 1}$ was sufficient to produce models with both good output fit and small parameter covariance, in the presence of an additive output white noise intensity corresponding to that of Figure 4a.

V. RESULTS

A. Example Patient Model

In this section we demonstrate the proposed method, using a realistic example from anesthesia control. We will use a PK model with parameters computed from demographic parameters, according to a formula by Schnider [10]. For our example we will assume that the patient is male, 30 years old, weighs 70 kg, and is 174 cm tall, resulting in the PK model:

$$\dot{x} = \frac{1}{60} \begin{bmatrix} -CL_1 + CL_2 + CL_3 & CL_2 & CL_3 \\ V_1 & V_1 & V_1 \\ CL_2 & CL_2 & 0 \\ V_2 & V_2 & 0 \\ CL_3 & 0 & CL_3 \\ V_2 & 0 & V_3 \end{bmatrix} x + \frac{1}{60^2} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u,$$

$$C_p = [1 \ 0 \ 0] x,$$

(7)

with clearances $CL = [1.68 \ 1.82 \ 0.84]^T$ l min$^{-1}$, and (virtual) compartment volumes $V = [4.27 \ 27.50 \ 238]^T$ l. The input $u$ (propofol infusion rate) is of unit mg h$^{-1}$ and the output $C_p$ (plasma concentration) is of unit $\mu g\cdot ml^{-1}$. The states are the per compartment drug concentrations. The PK model (7) is combined with a PD consisting of the first order system

$$G_{C_p,C_e} = \frac{k_{ce}}{s + k_{ce}}$$

(8)

relating the plasma concentration $C_p$ to the effect site concentration $C_e$, with $k_{ce} = 0.46/60$ s$^{-1}$, and the static output nonlinearity

$$E = \frac{v^\gamma}{v^\gamma + 1},$$

(9)

where $v = C_e/C_{e,50}$, $C_{e,50} = 1.8 \ \mu g\cdot ml^{-1}$ and $\gamma = 5.8$ (all being clinically relevant values). Apart from the patient model, the NeuroSense monitor [14], used to measure the clinical effect, has low-pass LTI dynamics:

$$G_{NS}(s) = \frac{1}{(8s + 1)^2},$$

(10)

relating the measurement $y$ to the clinical effect $E$. The typical operating point lies close to 50% of clinical effect, i.e., $C_e = C_{e,50}$ and $E = 0.5$ (corresponding to stationary input $u_0 = 181 \ \text{mg} \cdot \text{h}^{-1}$ and output $y = 0.5$). Linearizing the combined PKPD model and monitor dynamics (7)-(10) around this point, we obtain the transfer function

$$P_{in} = G_{u,C_p} \cdot G_{C_p,C_e} \cdot \frac{\gamma}{4C_{e,50}} \cdot G_{NS},$$

(11)

where $G_{u,C_p}$ is the transfer function corresponding to (7).

B. Experiment and Identification

The outcome of the proposed experiment is shown in Figure 4. The obtained FOTD model parameter vector is $\theta = [\tilde{b} \ \tilde{a} \ \tilde{L}]^T = [\bar{K}/T \ 1/T \ \bar{L}] = [1.42 \cdot 10^{-5} \ 2.82 \cdot 10^{3} \ 55]^T$, corresponding to

$$\check{P}(s) = \frac{K}{s^2 + 1 \ e^{-sL}} = \frac{5.04 \cdot 10^{-3}}{354s + 1} \ e^{-55s},$$

(12)

Over-estimation of the delay results from approximating the high order dynamics (11) by an FOTD system. The natural logarithm of the parameter covariance estimate is

$$\log(R_\theta) = \begin{bmatrix} -28.2 & -21.9 & -13.0 \\ -21.9 & -14.8 & -6.0 \\ -13.0 & -6.0 & 3.3 \end{bmatrix},$$

(13)

resulting in relative parameter standard deviations

$$\delta_\theta = \sqrt{\text{diag}(R_\theta)/\theta} = [5.2 \ 21 \ 9.5]^T 10^{-2},$$

where division is element-wise. To get an additional sense of model quality, (12) can be compared with the FOTD model obtained by balanced reduction of the delay-free part of (11), while keeping the delay unchanged:

$$P_{bal}(s) = \frac{5.14 \cdot 10^{-3}}{549s + 1} \ e^{-16.5s}.$$  

(14)

1In some literature, (9) has two additional calibration parameters $E_0$ and $E_{\text{max}}$. The version presented here corresponds to the (calibrated) case of $E_0 = 0$ and $E_{\text{max}} = 1$. 

\[\text{In some literature, (9) has two additional calibration parameters $E_0$ and $E_{\text{max}}$. The version presented here corresponds to the (calibrated) case of $E_0 = 0$ and $E_{\text{max}} = 1$.} \]
VI. DISCUSSION

This paper has presented a novel combination of a relay based experiment and an output error identification scheme. The main strengths lie in the short experiment duration and excitation at a phase shift relevant to control (inherent to relay methods). The method works reliably in the presence of noise and provides an estimate of the parameter covariance, in addition to nominal values.

In this work the method was demonstrated for identification of FOTD models. However, given sufficient excitation and initialization, it works equally well for higher order models.

It can also be noted that the described method allows for identification of the patient dynamics, with the monitor dynamics (10) excluded. This is enabled by applying (10) to $u$, prior to the identification.

REFERENCES