

Parameter identification of a tumor model using artificial neural networks

1st Melánia Puskás

Research and Innovation Center of Óbuda University
Óbuda University
Budapest, Hungary
puskas.melania@uni-obuda.hu

2nd Dániel András Drexler

Research and Innovation Center of Óbuda University
Óbuda University
Budapest, Hungary
drexler.daniel@uni-obuda.hu

Abstract—Mathematical models of tumor growth and the effect of therapy is fundamental for personalizing and optimizing anticancer therapies. The aim of the research is to provide a good estimation of personalized tumor model parameters based on measurements carried out on the patient. We use in silico experiments to create a large set of training data in a span that covers real-life scenarios. The data are used to train neural networks which provide a good initial guess for the model parameters. The estimated parameters can be used as initial estimations for more sophisticated, but local identification algorithms.

Index Terms—tumor model, personalizing therapy, neural networks, identification

I. INTRODUCTION

Optimization of cancer therapies based on mathematical models may hide great benefits, like minimizing drug dosage, side effects and costs, and preventing drug resistance [1]–[4]. Model-based optimization techniques can vary on a great scale using classical results from optimization and control engineering [5]–[7], while a few of them involves impulsive control, which models the treatments carried out using injections [4], [8]–[10]. Here we suppose that the treatment is done using injections, thus the tumor model is impulsive [11], [12].

There are numerous mathematical models of tumor growth dynamics [13], [14], some model tumor growth and supportive vasculature growth [15], some model the growth of tumor cells, healthy cells and immune cells [16], and some model tumor cell dynamics involving drug resistance [1]. We use a mathematical model which describes fundamental physiological phenomena of the tumor and the drug, like proliferation, necrosis, dead cell washout, pharmacodynamics and pharmacokinetics of the drug [2], [17]. The used tumor model is described in Subsection II-A.

Identification of the model parameters is crucial for efficient therapy optimization. Although, some parameters may not be identifiable from the measurements or their sensitivity can

be extremely low [18], which makes estimation of its value almost impossible. If we could build a database of parameter sets for specific patient populations, these could be used to generate offline optimized therapies, thus therapy optimization would not require great amount of real-time measurements during therapies. However, this has to be preceded by a great number of experiments on the specific patient population, and identification algorithms that can handle the constraint that interpatient variability should be kept low. Mixed-effect models can handle this problem [2], [19], [20], although they require a good initial guess for the parameters.

In order to provide a good estimation of the model parameters based on measurements, we use neural networks with two-layered, feed-forward architecture detailed in Subsection II-B. We solve two specific problems, first, we create a neural network to estimate the pharmacokinetic parameters of the drug in the patient based on measurements of blood serum level carried out right after the injection of the drug every nine minutes. We will call this network the PKNN (pharmacokinetics neural network). Second, we create a neural network to estimate the tumor growth parameters and the pharmacodynamics parameters of the drug, based on tumor volume and blood serum level measurements carried out in every two days. We will call this network the TDNN (tumor dynamics neural network). We generate a large set of training data for these neural network to cover real-life scenarios in Section III.

The training of the PKNN results in reliable estimator for most of the training data which can provide accurate parameters values for most of the cases as it will be shown in Subsection IV-A. The training of the TDNN shows diverse results which are in accordance with the results of the identification analysis by Siket et al [18]. The resulting neural networks can be applied for therapy optimization, parameters set collection on specific patient population and for following the trends in the inpatient variability.

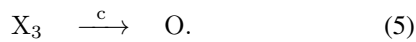
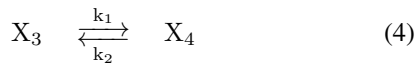
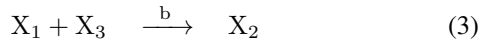
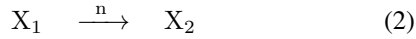
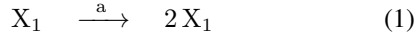
II. PRELIMINARIES

A. Tumor model

The tumor model used in the research is composed of ordinary differential equations that model basic physiological phenomena. The meaning of the terms in the equation can be

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 679681). Project no. 2019-1.3.1-KK-2019-00007. has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the 2019-1.3.1-KK funding scheme. The present work has also been supported by the Hungarian National Research, Development and Innovation Office (2018-2.1.11-T ÉT-SI-2018-00007 and SNN 125739).

visualized to medical experts by using formal reaction kinetics analogy [2]. Let X_1 , X_2 , X_3 and X_4 denote the fictive species representing the living tumor volume, dead tumor volume, drug level in central compartment and drug level in peripheral compartment, respectively. Then the modeled phenomena are described as



where (1) is tumor proliferation with rate a , (2) is tumor necrosis with rate n , (3) is the effect of the drug (pharmacodynamics) with maximal effect rate b and median effective dose ED_{50} (the latter is not depicted in (3)), (4) is the drug transport between the central and peripheral compartment, while (5) is the depletion of the drug.

These formal reaction steps are considered with mass-action kinetics to get the ordinary differential equations of the tumor model [21], [22], except for (3), which is partly considered with Michaelis-Menten kinetics with constant ED_{50} to model saturation of the drug effect. The resulting ordinary differential equations are

$$\dot{x}_1 = (a - n)x_1 - b \frac{x_1 x_3}{ED_{50} + x_3} \quad (6)$$

$$\dot{x}_2 = nx_1 + b \frac{x_1 x_3}{ED_{50} + x_3} - wx_2 \quad (7)$$

$$\dot{x}_3 = -(c + k_1)x_3 + k_2 x_4 - b_k \frac{x_1 x_3}{ED_{50} + x_3} \quad (8)$$

$$\dot{x}_4 = k_1 x_3 - k_2 x_4. \quad (9)$$

In these equations, x_1 is the time function of the living tumor volume [mm^3], x_2 is the time function of the dead tumor volume [mm^3], x_3 is the time function of the drug serum level in the central compartment (blood) [ng/ml], and x_4 is the time function of the drug serum level in the peripheral compartment (tissues) [ng/ml].

Previous identifications results showed that the parameter b_k is close to zero [2], [17], so we will use $b_k = 0$ in the remaining of the paper.

The model is impulsive, thus the input is impulsive effect on x_3 , which represents injections into the central compartment. Let t_k , $k = 1, 2, \dots$ be the time of injections, while u_k , $k = 1, 2, \dots$ is the dosage of injections in mg/kg . Then the function x_3 has a discontinuity of the first kind at time instants t_k , modeled as

$$x_3(t_k^+) = x_3(t_k^-) + \kappa u_k, \quad (10)$$

where κ is a patient dependent constant which transforms the dosage [mg/kg] into serum level [ng/ml].

The model parameters have been identified using mixed-effect model [2], [17], although this method was based on the assumption that the parameters are constant. However,

this assumption greatly limits the modeling power of (6)-(9). A potential application of the neural network is parameter estimation with a sliding window with relatively small time intervals, which could be used to track the parameter changes in the patient.

B. Neural network

We use a two-layer feed-forward neural network, with one hidden layer and one output layer. The output layer contains linear neurons, i.e., the outputs are weighted sums of the outputs of the neurons in the first layer and a bias. The output of the r th neuron of the first layer are calculated as

$$y_r^{(1)}(\mathbf{x}) = \varphi \left(\sum_{i=1}^{N_i} w_{r,i}^{(1)} x_i(p) + \underbrace{w_{j,N_i+1}^{(1)}}_{\text{offset}} \right), \quad (11)$$

where $r = 1, 2, \dots, R$ and R is the number of neurons in the hidden layer, the superscript (1) refers to layer 1, $\mathbf{x}$ is the input of the neural network, x_i is the i th input of the neural network, where $i = 1, 2, \dots, N_i$ and N_i is the number of inputs, $w_{r,i}^{(1)}$ is the weight of the i th input of the r th neuron in layer 1, and $w_{r,N_i+1}^{(1)}$ is the weight of the bias of the r th neuron. The activation function of the neurons in the hidden layer is the sigmoid function

$$\varphi(\epsilon) = \frac{2}{1 + e^{-2\epsilon}} - 1. \quad (12)$$

The m th output of the neural network is

$$y_m^{(2)}(\mathbf{x}) = \sum_{r=1}^R w_{m,r}^{(2)} y_r^{(1)}(\mathbf{x}) + \underbrace{w_{m,R+1}^{(2)}}_{\text{offset}}, \quad (13)$$

where the superscript (2) refers to layer 2, $w_{m,r}^{(2)}$ is the weight of the m th neuron and r th input in the second layer, and $w_{m,R+1}^{(2)}$ is the weight of the bias for the m th neuron in the output layer.

The neural network is trained with supervised learning utilizing error backpropagation using a Levenberg-Marquardt algorithm [23]. The training data are generated with in silico experiments from a large parameter set.

III. TRAINING DATA

We generated training data sets for both the PKNN and TDNN, the data sets consisted of 100 variations for 1000 different patients, i.e., 100000 training input-output pair were generated. Due to the different nature of the experimental setups where the PKNN and TDNN are used, we have created the training data in a different way for these networks.

The inputs of the PKNN are measurements carried out on blood samples taken from the patient every nine minutes, the first sample is taken five minutes after the injection of the drug. We suppose that there is no drug present in the patient before the injection, i.e., the initial conditions for the two compartment model (8)-(9) is zero in the in silico data generation process. We have created training sets for 5, 9 and 18 number of measurements, the inputs of the PKNN are the

measurements and the injected dosage, thus the number of inputs of the PKNN were 6, 10, and 19, respectively. For practical reasons, fewer measurements are desirable. After evaluating the training results, we concluded that the PKNN which operated with five measurement points has similar performance as the other two PKNNs which operated with 9 and 18 measurements, thus we finally chose the PKNN with five measurements, which is discussed in Subsection IV-A.

The inputs of the TDNN are measurements of the total tumor volume and blood serum levels acquired from blood samples, the measurements are done every two days. The initial conditions are supposed to be the same for all the patients, with $x_1(0) = 200 \text{ mm}^3$, which is the tumor volume where treatment begins, $x_2(0) = 0 \text{ mm}^3$, and with zero drug levels in the compartments, since in two days the drug either depletes from the body or the serum level becomes extremely small. We have created training sets for 6, 11 and 16 number of measurements, the inputs of the TDNN are the total tumor volume measurements and the measured drug serum levels, thus the number of inputs of the TDNN were 12, 22, and 32, respectively. For practical reasons, fewer measurements are desirable here as well. After evaluating the training results, we concluded that the TDNN which operated with six measurement points has similar performance as the other two TDNNs which operated with 11 and 16 measurements, thus we finally chose the TDNN with six measurements, which is discussed in Subsection IV-B.

IV. TRAINING RESULTS

A. Pharmacokinetic Neural Network

The error histogram after the training of the PKNN operating with five measurements is in Fig. 1. The diagram shows that the error for most of the training, validation and test data are close to zero, which implies that the trained PKNN can estimate the pharmacokinetic parameters of the drug inside the patient with five measurements taken after the injections.

Figure 2 shows the relative estimation errors of the different pharmacokinetic parameters for all the data. The estimation error for most of the parameters is close to zero, depicted with blue rectangles, however, there are some outliers for parameter a which have extremely large relative estimation error.

Figure 3 shows the relative estimation errors in the range $\pm 20\%$. The percentages over the parameters on the top of the figure show the how many of the total data from the data set falls into this range. The results show that the parameters can be estimated within the 20% range for most of the data. The worst parameter in Fig. 3 is the parameter κ , however Fig. 2 shows that the estimation results examined on all the data are the smallest for this parameter.

B. Tumor Dynamics Neural Network

The result of the parameter estimation for the trained TDNN is shown in Fig. 4. The relative estimation errors are shown for each parameters for all the data. There are extreme large outliers for the necrotic rate n and the median effective dose ED_{50} . The proliferation rate a , maximum drug effect rate b

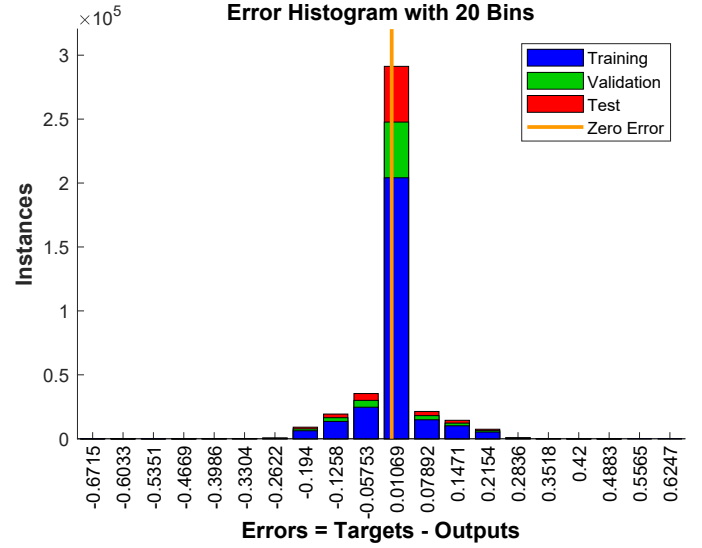


Fig. 1. The error histogram after training the PKNN with 100000 samples.

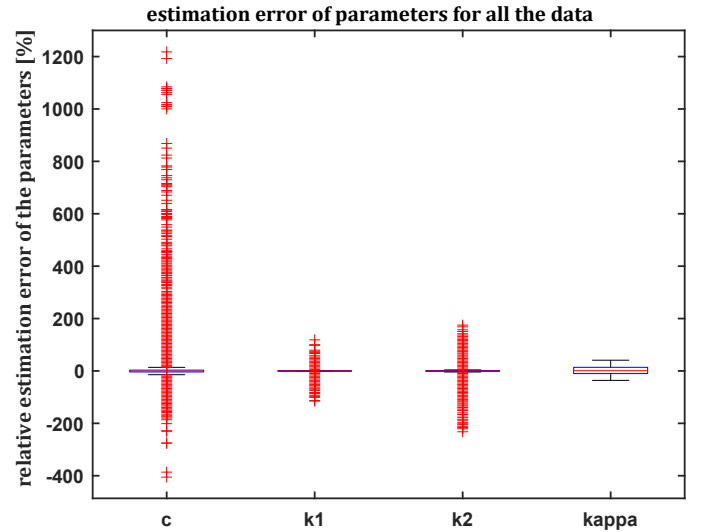


Fig. 2. The relative estimation error of the PKNN on all the 10000 samples, 50% of the data are in the blue rectangles, the red crosses are outliers.

and the washout rate w can be estimated with good accuracy, thus the TDNN can be used to give a good initial guess for the latter three parameters.

Figure 5 shows the relative estimation errors in the range $\pm 20\%$. The estimation error for parameter a stays in this range for all the data, thus this parameter can be estimated reliably. A large portion of the data falls into this range for parameters w and b , although only a small portion for the parameters n and ED_{50} , thus these parameters can not be estimated reliably. The sensitivity analysis carried out in Siket et al. also showed that the effect of the local change of the parameter ED_{50} has the least effect on the output [18]. This is also justified by the training results of the TDNN.

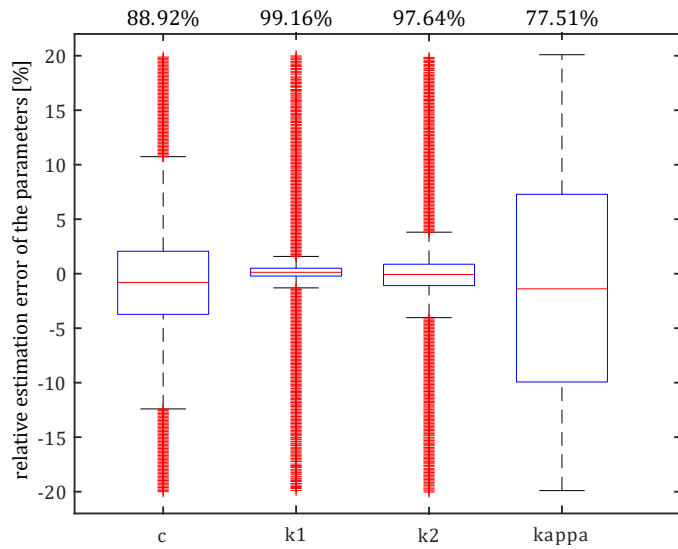


Fig. 3. The relative estimation error of the PKNN in the range of $\pm 20\%$, indicating the percentage of the samples in this range (on top of the figure for each parameter), 50% of the data are in the blue rectangles, the red crosses are outliers, the red segments are mean values.

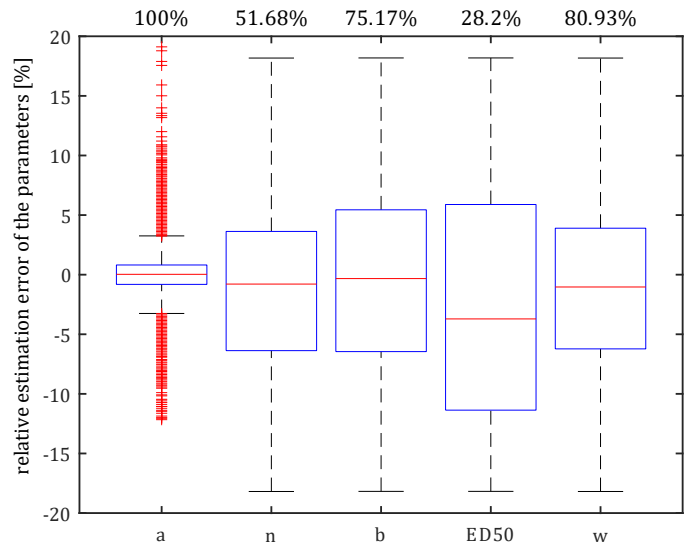


Fig. 5. The relative estimation error of the TDNN in the range of $\pm 20\%$, indicating the percentage of the samples in this range (on top of the figure for each parameter), 50% of the data are in the blue rectangles, the red crosses are outliers, the red segments are mean values.

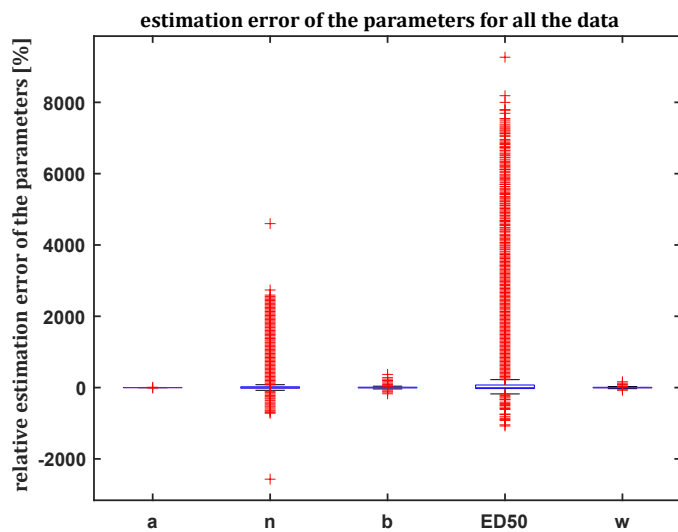


Fig. 4. The relative estimation error of the TDNN on all the 10000 samples, 50% of the data are in the blue rectangles, the red crosses are outliers.

V. CONCLUSION

Two different neural networks were trained with large data sets acquired from in silico experiments. The data sets span a large set of realistic model parameters. The results show that the pharmacokinetic parameters can be estimated reliably except for a very few cases. The estimation results for the tumor dynamics parameters show that the proliferation rate of the living cells, the washout rate of the dead cells and the inhibitory rate with maximum drug effect can be estimated well with the neural network, however, the estimation of the median effective dose parameter and the necrotic rate is less reliable.

REFERENCES

- [1] J. M. Greene, J. L. Gevertz, and E. D. Sontag, "Mathematical approach to differentiate spontaneous and induced evolution to drug resistance during cancer treatment," *JCO Clinical Cancer Informatics*, no. 3, pp. 1–20, 2019.
- [2] D. A. Drexler, T. Ferenci, A. Lovrics, and L. Kovács, "Tumor dynamics modeling based on formal reaction kinetics," *Acta Polytechnica Hungarica*, vol. 16, pp. 31–44, 2019.
- [3] A. R. Akhmetzhanov, J. W. Kim, R. Sullivan, R. A. Beckman, P. Tamayo, and C.-H. Yeang, "Modelling bistable tumour population dynamics to design effective treatment strategies," *Journal of theoretical biology*, vol. 474, pp. 88–102, 2019.
- [4] F. Cacace, V. Cusimano, and P. Palumbo, "Optimal impulsive control with application to antiangiogenic tumor therapy," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 1, pp. 106–117, Jan 2020.
- [5] J. Klamka, H. Maurer, and A. Swierniak, "Local controllability and optimal control for a model of combined anticancer therapy with control delays," *Mathematical Biosciences and Engineering*, vol. 14, no. 1, pp. 195–216, 2017.
- [6] F. S. Lobato, V. S. Machado, and V. Steffen, "Determination of an optimal control strategy for drug administration in tumor treatment using multi-objective optimization differential evolution," *Computer Methods and Programs in Biomedicine*, vol. 131, pp. 51 – 61, 2016.
- [7] D. A. Drexler, J. Sági, and L. Kovács, " H_∞ control of nonlinear systems with positive input with application to antiangiogenic therapy," in *Proceedings of the 9th IFAC Symposium on Robust Control Design ROCOND 2018*, 2018, pp. 146 – 151.
- [8] H.-P. Ren, Y. Yang, M. S. Baptista, and C. Grebogi, "Tumour chemotherapy strategy based on impulse control theory," *Philosophical Transactions Mathematical Physical & Engineering Sciences*, vol. 375, no. 2088, 2017.
- [9] D. A. Drexler and L. Kovács, "Optimization of impulsive discrete-time tumor chemotherapy," in *Proceedings of the 2019 IEEE 1st International Conference on Societal Automation*, 2019.
- [10] B. Péceli, D. A. Drexler, and L. Kovács, "Optimal scheduling of low-dose metronomic chemotherapy: an in-silico analysis," in *2020 IEEE 15th International Conference of System of Systems Engineering (SoSE)*, 2020, pp. 411–416.
- [11] N. A. Perestyuk, V. A. Plotnikov, A. M. Samoilenko, and N. V. Skripnik, *Differential Equations with Impulse Effects*. Berlin, Germany: De Gruyter, 2011.
- [12] T. Yang, *Impulsive Control Theory*. Berlin, Germany: Springer, 2001.

- [13] A. M. Jarrett, E. A. B. F. Lima, D. A. Hormuth, M. T. McKenna, X. Feng, D. A. Ekrut, A. C. M. Resende, A. Brock, and T. E. Yankeelov, "Mathematical models of tumor cell proliferation: A review of the literature," *Expert Review of Anticancer Therapy*, vol. 18, no. 12, pp. 1271–1286, 2018.
- [14] P. M. Altrock, L. L. Liu, and F. Michor, "The mathematics of cancer: integrating quantitative models," *Nature Reviews. Cancer*, vol. 15, no. 12, pp. 730–745, 2015.
- [15] P. Hahnfeldt, D. Panigrahy, J. Folkman, and L. Hlatky, "Tumor development under angiogenic signaling: A dynamical theory of tumor growth, treatment response, and postvascular dormancy," *Cancer Research*, vol. 59, pp. 4770–4775, 1999.
- [16] P. A. Valle, L. N. Coria, and Y. Salazar, "Tumor Clearance Analysis on a Cancer Chemo-Immunotherapy Mathematical Model," *Bulletin of Mathematical Biology*, vol. 81, no. 10, pp. 4144–4173, Oct. 2019, wOS:000487900700018.
- [17] D. A. Drexler, T. Ferenci, A. Füredi, G. Szakács, and L. Kovács, "Experimental data-driven tumor modeling for chemotherapy," in *Proceedings of the 21st International World Congress of the International Federation of Automatic Control*, 2020, accepted for publication.
- [18] M. Siket, G. Eigner, and L. Kovács, "Sensitivity and identifiability analysis of a third-order tumor growth model," in *2020 IEEE 15th International Conference of System of Systems Engineering (SoSE)*, 2020, pp. 417–422.
- [19] M. Fidler, Y. Xiong, R. Schoemaker, J. Wilkins, M. Trame, T. Post, and W. Wang, *nlmixr: Nonlinear Mixed Effects Models in Population Pharmacokinetics and Pharmacodynamics*, 2018, r package version 1.0.0-7. [Online]. Available: <https://CRAN.R-project.org/package=nlmixr>
- [20] J. Owen and J. Fiedler-Kelly, *Introduction to Population Pharmacokinetic / Pharmacodynamic Analysis with Nonlinear Mixed Effects Models*. Wiley, 2014.
- [21] P. Erdi and J. Tóth, *Mathematical Models of Chemical Reactions. Theory and Applications of Deterministic and Stochastic Models*. Princeton, New Jersey: Princeton University Press, 1989.
- [22] J. Tóth, A. L. Nagy, and D. Papp, *Reaction kinetics: exercises, programs and theorems*. Springer, 2018.
- [23] H. Yu and B. Wilamowski, *Levenberg–Marquardt Training*, 02 2011.

