LOTUS-LIKE PLOT, ANNULUS & DISK: ON MODEL-BASED DESIGN OF PHOTOBLEACHING EXPERIMENTS

Š. Papáček♠, J. Jablonský♠, D. Štys♠, K. Petera♢, C. Matonoha♣

♠Institute of Complex Systems, FFPW, University of South Bohemia, Zamek 136, 373 33 Nove Hrady, Czechia
♢Czech Technical University in Prague, Faculty of Mechanical Engineering, Technicka 4, 166 07 Prague 6, CR
♣Institute of Computer Science, Academy of Sciences of the Czech Republic, 182 07 Prague 8, Czech Republic

Abstract

Our work is focused on the so-called model-based design of experiments (MBDOE) and shows practical applications in biological research. Although the MBDOE method is model independent, we restrict ourselves to the parameter estimation problem for a reaction-diffusion PDE system in form of the Initial Boundary Value Problem (IBVP). The presented case study is focused on the FRAP (Fluorescence Recovery After Photobleaching) experimental technique applied to study the mobility of photosynthetic protein complexes within the Institute of Microbiology CAS, Třeboň, CZ, and we reached rather surprising results, e.g., (i) *Ill-posedness* of the inverse problem when the diffusivity can vary in time, (ii) *Lotus-like plot* for the optimal data space selection, (iii) *Disk vs. Annulus* , alternating winners of the competition for the best bleach topology.

1 STATE-OF-THE-ART OF FRAP DATA PROCESSING

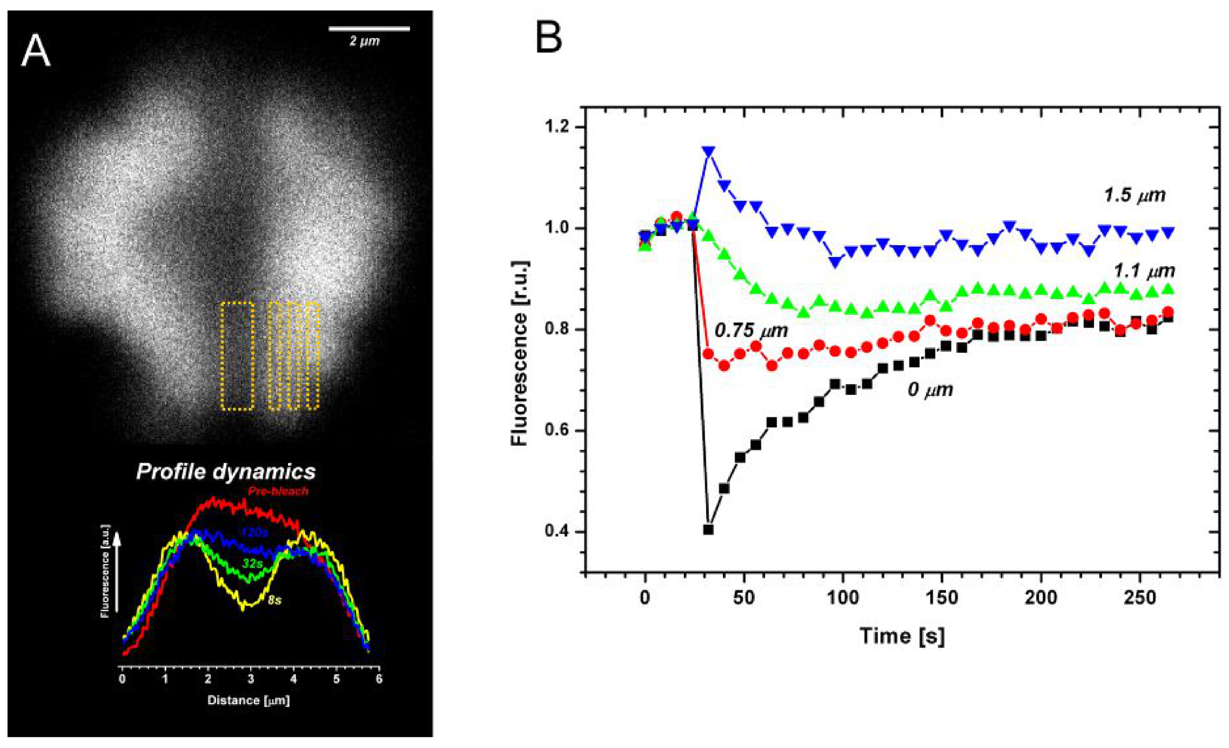

Fig. 1 **A - Upper left:** Image of red algae *P. Cruentum* acquired by the confocal laser scanning microscopy (CLSM) 8 s after bleaching. Four regions of interest (**ROI**) are labeled as yellow rectangles. **A - Lower left:** Recovery dynamics in time-sequence of one-dimensional **bleach profiles** perpendicular to the bleach stripe. The fluorescence in central ROI is recovered with the growing time. **B:** Time series of the space-averaged fluorescence signal over 4 different ROI. The lowest curve corresponding to the bleached ROI is the so-called **FRAP recovery curve**, Kaňa *et al.* (2014).

2 MATHEMATICAL MODEL

Initial Boundary Value Problem (IBVP)

Let us have $u = (u_1, ..., u_{q_c})^T$ vector of concentrations of q_c a priori not fixed interacting components, D a diagonal matrix of diffusion coefficients $(D_k)_{k=1}^{q_c}$, and K a matrix of reaction rates, then the reaction-diffusion PDE is

$$\frac{\partial u(x,t)}{\partial t} = D\Delta u(x,t) - Ku(x,t),\ x \in \Omega,\ t \in [0,T], \tag{1}$$

with **initial conditions** $u_k(x,0) = u_{k0}\phi(x),\ k \in \{1,\dots,q_c\}$, $\phi(x)$ is an initial shape.

Boundary conditions could be of Dirichlet or Neumann type, e.g., $u(x,t) = 0,$ or $\frac{\partial u(x,t)}{\partial n} = 0,$ on $\partial\Omega \times [0,T]$. The **measured fluorescent signal** y (on spatio-temporal grid) is proportional to the sum of concentration profiles $\sum_{k=1}^{q_c} u_k(x_i,t_j)$.

Model-parameter estimation – Inverse (ill-posed) problem !?

We define the forward map (also called the parameter-to-data map)

$$F : \mathcal{R}^q \to \mathcal{R}^{N_{\text{data}}}, \quad F(p) = (y(x_i,t_i))_{i=1}^{N_{\text{data}}},$$

where the total number of parameters to identify is $q = q_c + q_r$, i.e., $p \in \mathcal{R}^q$.

Our regression model for the parameter vector p, is

$$F(p) = \text{data}, \tag{2}$$

where the data are usually contaminated with additive white noise

$$\text{data} = F(p_T) + e = (y(x_i,t_i))_{i=1}^{N_{\text{data}}} + (e_i)_{i=1}^{N_{\text{data}}}, \tag{3}$$

i.e., $e(t_j) = \mathcal{N}(0,\sigma^2)$, for each time instant t_j, and $p_T \in \mathcal{R}^q$ are true coefficients.

The aim of the parameter estimation is to find $p \in \mathcal{R}^q$ such that (4) is satisfied:

$$\|F(p_c) - \text{data}\|^2 = \min_p \|F(p) - \text{data}\|^2. \tag{4}$$

This is usually **ill-posed problem**, thus a regularization technique has to be employed, see e.g., Engl, Hanke, Neubauer (1996), Papáček *et al.* (2013).

3 OPTIMIZATION OF AN EXPERIMENT DESIGN

Sensitivity and Fisher information matrix

First, we require the Fréchet-derivative $F'[p] \in \mathcal{R}^{N_{\text{data}} \times q}$ of the forward map F:

$$F'[p] = \frac{\partial}{\partial p}F(p) = \left(\frac{\partial}{\partial p}y(x_1,t_1), \quad \dots, \quad \frac{\partial}{\partial p}y(x_{N_{\text{data}}}, t_{N_{\text{data}}})\right)^T.$$

A corresponding quantity (a "global sensitivity") is the **Fisher information matrix**

$$M = F'[p]^T F'[p] \in \mathcal{R}^{q \times q}. \tag{5}$$

In order to discriminate between *relevant* and *irrelevant* data, we compute $(\frac{\partial}{\partial p}y(x,t))$ and select the data regions by ordering all data points according to their "local sensitivities" $(\frac{\partial}{\partial p}y(x_1,t_1))^2 \le (\frac{\partial}{\partial p}y(x_2,t_2))^2 \le$ ... Then we select the reduced data from the region of points $(x_i,t_i),\ i = 1,\dots,\bar{N}_{\text{data}}$ using this ordering. I.e., we take the most sensitive data points as relevant data (for some fixed value of η). The resulting visualization is the *lotus-like plot*, cf. Fig. 2, where the gray regions correspond to regions where (for a previously chosen threshold η) the relevant data $(\bar{N}_{\text{data}})$ are taken, Kindermann and Papáček (2015).

The key relation is:

$$(\bar{p}_c - p_T)^2 \sum_{i=1}^{N_{\text{data}}} \left[\frac{\partial}{\partial p}y(x_i,t_i)\right]^2 \le (1+\eta)\frac{\text{res}^2(\bar{p}_c)}{\bar{N}_{\text{data}} - 1} f_{1,N_{\text{data}}-1}(\alpha) \approx \sigma^2. \tag{6}$$

Lemma 1: Let $\bar{N}_{\text{data}}$ be the number of reduced (relevant) data points and suppose that η is chosen according to $\eta \ge \frac{\sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}}(\frac{\partial}{\partial p}y(x_i,t_i))^2}{\sum_{i=1}^{\bar{N}_{\text{data}}}(\frac{\partial}{\partial p}y(x_i,t_i))^2}$. Then using reduced data, we find a confidence interval (for $\bar{p}_c$) of the form (6). The ratio $\frac{\bar{N}_{\text{data}}}{N_{\text{data}}}$ indicates the fraction of "η-relevant" data points taken over all data points. The factor $1+\eta$ represents the enlargement of the confidence interval compared to p_c.

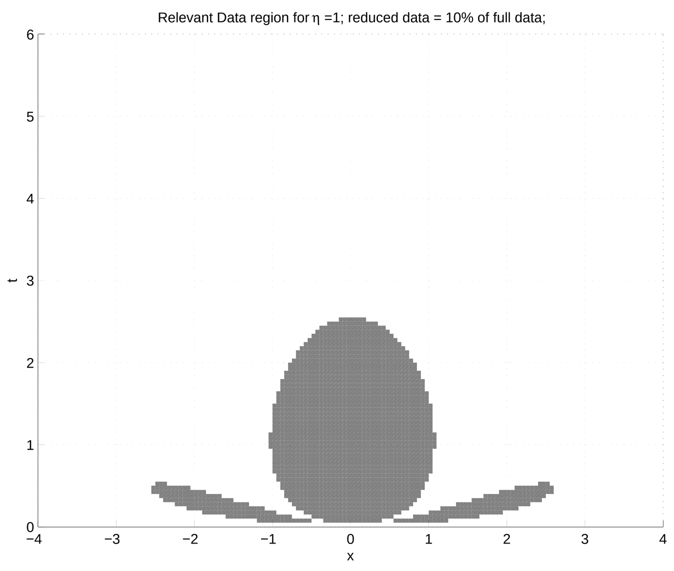

Fig. 2a: Ratio $\frac{\bar{N}_{\text{data}}}{N_{\text{data}}} = 0.1$ was determined for $\eta = 1$, i.e., confidence interval for $\bar{p}_c$ compared to p_c is twofold, but only 10% of data is taken.

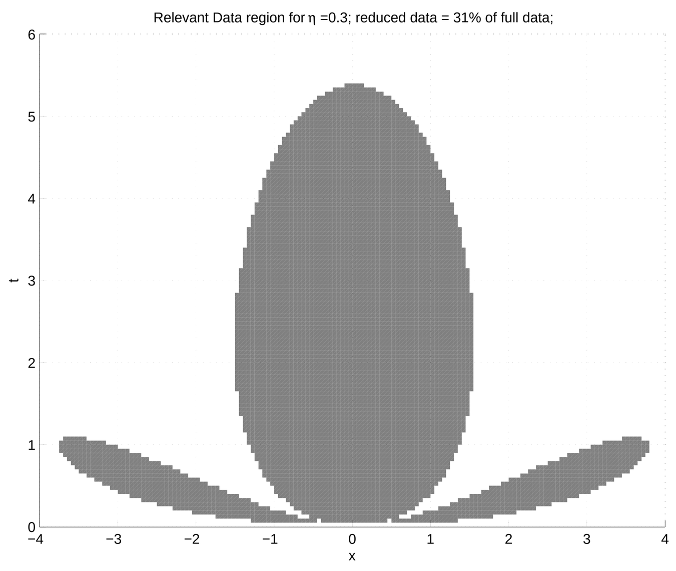

Fig. 2b: A lotus-like plot for the ratio $\frac{\bar{N}_{\text{data}}}{N_{\text{data}}} = 0.31$, corresponding to $\eta = 0.3$. For the signification of lateral petals, see Papáček *et al.* (2015).

When the bleach topology is subject to optimization...

For radially symmetric problem (R is the size of monitored ROI, T is the experiment duration), the winner of competition for the best bleach topology depends on $\frac{R^2}{4TD}$, see Fig. 3.

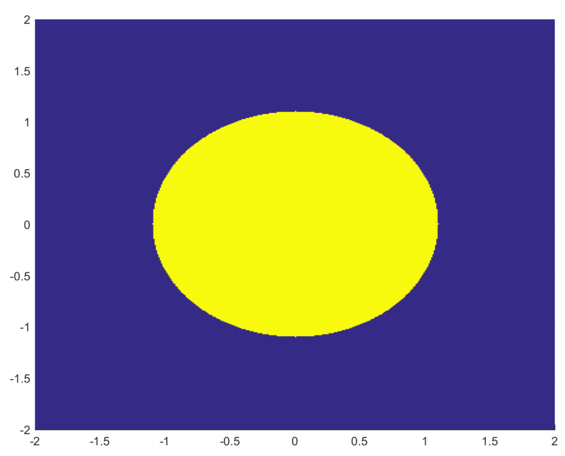

Fig. 3a: Disk as optimal bleach shape for the fast diffusion, i.e., low diffusion time t_{char}: $\frac{R^2}{4TD} = \frac{t_{char}}{T} < 1.8$.

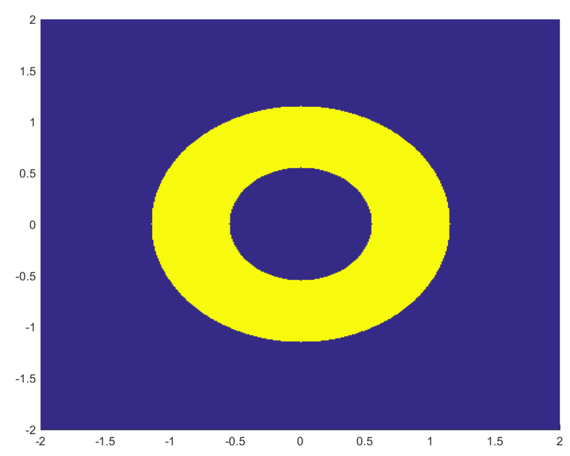

Fig. 3b: Optimal bleach shape for higher diffusion time $1.8 < \frac{R^2}{4TD} < 6$. For more details, Kindermann and Papáček (2016).

4 CONCLUSION

Our study started with the question: Could the MBDOE method enhance the precision of parameter estimates? Then the problem of the data space selection and the optimal initial bleach shape for the FRAP model parameter identification was formulated. As an optimality criterion, we choose to maximize the sensitivity measure M, Eq. (5), in order to have the expected error minimal. We studied the optimization problem with the fixed bleach depth, finding the analytical expressions for the sensitivity measure M, allowing to find the relevant data space and compare different initial bleach shapes, as well. While the first task was visualized via the *lotus-like plot*, the latter revealed even more surprising result: For small values of the scaled inverse diffusion coefficient (or low values of $\frac{t_{char}}{T}$), the *disk* is the optimal shape, while for higher values, shapes with more and more components (i.e. *annuli-type shapes*) become optimal (and realizable, cf. Blumenthal *et al.* (2015)), leading to a significant improvement in the parameter precision, Kindermann and Papáček (2016).

Acknowledgement

This work was supported by the Ministry of Education, Youth and Sport of the Czech Republic , projects CENAKVA (No. CZ.1.05/2.1.00/01.0024) and CENAKVA II (No.LO1205 under the NPU I program), by the long-term strategic development financing of the Institute of Computer Science (RVO:67985807) and by the Austrian agency for international mobility and cooperation in education, science and research within the programme "Aktion Oesterreich-Tschechien". SP would like to thank to Prof. Stefan Kindermann from the Industrial Mathematics Institute, Johannes Kepler University of Linz, for valuable discussions and hard work on "FRAP project".

References

Blumenthal, D., Goldstien, L., Edidin, M., Gheber, L.A. (2015) *Universal approach to FRAP analysis of arbitrary bleaching patterns.* Scientific Reports 5, 11655; doi: 10.1038/srep 11655.

Engl, H., Hanke, M., Neubauer, A. (1996) *Regularization of ill-posed problems.* Kluwer, Dortrecht.

Kaňa, R., Kotabová, E., Lukeš, M., Papáček, Š., Matonoha, C., Liu, L.N., Prášil, O., Mullineaux, C.W. (2014) *Phycobilisome mobility and its role in the regulation of light harvesting in red algae.* Plant Physiology 165(4), 1618–1631.

Kindermann, S., Papáček, Š. (2015) *On data space selection and data processing for parameter identification in a reaction-diffusion model based on FRAP experiments.* Abstract and Applied Analysis, Article ID 859849, 17 pp.

Kindermann, S., Papáček, Š. (2016) *On the optimality of the shape (and topology) of the initial conditions for FRAP experiments.* - submitted to PeerJ Computer Science.

Papáček, Š., Kaňa, R., Matonoha, C. (2013) *Estimation of diffusivity of phycobilisomes on thylakoid membrane based on spatio-temporal FRAP images.* Mathematical and Computer Modelling 57, 1907–1912.

Papáček, Š., Jablonský, J., Matonoha, C., Kaňa, R., Kindermann, S. (2015) *FRAP & FLIP: Two sides of the same coin?* Bioinformatics and Biomedical Engineering, Volume 9044 of the series Lecture Notes in Computer Science, 444–455.