

## Research Article

# On Data Space Selection and Data Processing for Parameter Identification in a Reaction-Diffusion Model Based on FRAP Experiments

Stefan Kindermann<sup>1</sup> and Štěpán Papáček<sup>2</sup>

<sup>1</sup>Industrial Mathematics Institute, Johannes Kepler University of Linz, Altenbergerstr 69, 4040 Linz, Austria

<sup>2</sup>Institute of Complex Systems, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, Faculty of Fisheries and Protection of Waters, University of South Bohemia-České Budějovice, Zámek 136, 373 33 Nové Hradky, Czech Republic

Correspondence should be addressed to Štěpán Papáček; [spapacek@frov.jcu.cz](mailto:spapacek@frov.jcu.cz)

Received 2 January 2015; Revised 15 May 2015; Accepted 18 May 2015

Academic Editor: Benito M. Chen-Charpentier

Copyright © 2015 S. Kindermann and Š. Papáček. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fluorescence recovery after photobleaching (FRAP) is a widely used measurement technique to determine the mobility of fluorescent molecules within living cells. While the experimental setup and protocol for FRAP experiments are usually fixed, data (pre)processing represents an important issue. The aim of this paper is twofold. First, we formulate and solve the problem of *relevant* FRAP data selection. The theoretical findings are illustrated by the comparison of the results of parameter identification when the full data set was used and the case when the *irrelevant data set* (data with negligible impact on the confidence interval of the estimated parameters) was removed from the data space. Second, we analyze and compare two approaches of FRAP data processing. Our proposition, surprisingly for the FRAP community, claims that the data set represented by the FRAP recovery curves in form of a time series (*integrated data approach* commonly used by the FRAP community) leads to a larger confidence interval compared to the *full (spatiotemporal) data approach*.

## 1. Introduction

The method of fluorescence recovery after photobleaching (FRAP) is based on the measurement of the change of fluorescence emitted by autofluorescent molecules or fluorescently tagged compounds (e.g., green fluorescence proteins (GFP) in a region of interest (ROI), usually a 2D Euclidean bounded domain, in response to a high-intensity laser pulse provided by confocal laser scanning microscopy (CLSM). The initial steady state is thus perturbed by an external stimulus, the so-called bleach. The bleach (or bleaching) causes an irreversible loss of fluorescence in the bleached area, apparently without any significant damage to intracellular structures. After bleaching, the observed recovery in fluorescence reflects the mobility of fluorescence compounds from the area outside the bleach [1].

To quantify the mobility of photosynthetic proteins of different microbial species is the main research interest of biologists. In Figure 1, we observe an example of FRAP data in form of a time series of both 2D and 1D fluorescence profiles reflecting the photosynthetic proteins mobility on a membrane [2, 3]. In the observed specimen (thylakoid membrane of the red algae *Porphyridium cruentum*, cf. Figure 1(a)), the experimentalists usually select a 2D rectangular ROI with the larger side perpendicular to the bleach strip in order to perform the averaging along the shorter axis. The resulting 1D fluorescence profiles (Gaussian-shaped signals in the central bleached region) are plotted on Figure 1(b). Based on spatiotemporal FRAP data, the diffusion constant is usually estimated by fitting a closed form model to the preprocessed FRAP data; see, for example, [4–9]. Since all

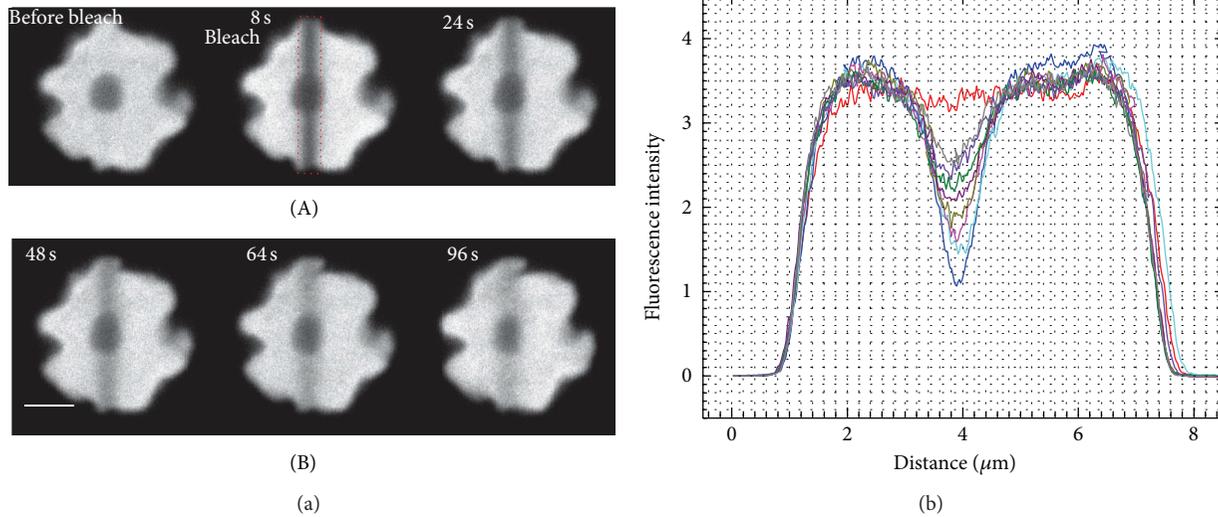


FIGURE 1: Experimental data from FRAP experiments [2, 3]. (a) Representative FRAP image sequence for a single cell of red algae *Porphyridium cruentum* for phycobilisome fluorescence. First, a fluorescence image before bleaching was detected ((A) before bleach), and then the phycobilisome fluorescence was bleached out across the middle of the cell in the vertical direction (red rectangle) and the postbleach images were recorded every 8 seconds (5 plots are displayed only). The total measurement time was 200 seconds; the length of the scale bar is  $3 \mu\text{m}$ . (b) One-dimensional fluorescence profiles. The abscissa represents the position along the axis perpendicular to the bleach strip. The ordinate quantifies the corresponding values of fluorescence (in arbitrary units) averaged along the axis parallel to the bleach strip. In the central region we see the step-wise recovery of the Gaussian-shaped signals, from the lowest values of the first postbleach profile to the highest prebleach (steady-state) values on the top.

closed form equation models need some simplified conditions to be fulfilled, for example, specific assumptions about initial and boundary conditions, they have limited accuracy. The simulation based models are more general although computationally more expensive, and they are using the raw data; the underlying process is modeled numerically; see, for example, PDE based model [10] or Monte Carlo simulations based model [11, 12], and model parameters are inferred within an optimization procedure.

Let us mention that CLSM allows the selection of a thin cross section of the sample by rejecting the information coming from the out-of-focus planes. However, the small energy level emitted by the fluorophore and the amplification performed by the photon detector is the cause of significant measurement noise. Therefore, FRAP images are in general very noisy with a small signal-to-noise ratio (SNR); that is, in order to get reliable results for the diffusion coefficient  $D$  (often modeled as time and space invariant), an adequate technique residing in regularization is mandatory [13]. The analysis of the ill-posedness of the parameter identification of reaction-diffusion models based on spatiotemporal FRAP images was treated in [14, 15] and previously also in [16]. Currently, our approach is implemented into the special software CA-FRAP based on the UFO (Universal Functional Optimization) system [17]. For the real FRAP data with the red algae *Porphyridium cruentum* we obtained a good agreement with reference values (the range of result  $10^{-14} \text{m}^2 \text{s}^{-1}$ , i.e.,  $10^{-2} \mu\text{m}^2 \text{s}^{-1}$ ) [3, 15, 18].

In this paper, we focus on another rather theoretical issue with great practical relevance residing in FRAP

data (pre)processing. The key concept relies on sensitivity analysis; confer [19]. Sensitivity analysis, performed to test the sensitivity of the parameter inference of particular modelling conditions, is well established. However, we are using sensitivity analysis to reduce data set size *prior* to conducting parameter inference; that is, we calculate the sensitivities and afterwards select the “relevant” data space where the sensitivity is “sufficiently high.” The importance of this point is notably raised when the cost of data acquisition (besides considering the computational issues related to the “irrelevant” data processing) cannot be neglected. As far as we know, this approach is novel in both the FRAP-related and statistic literature. Furthermore, based on the same sensitivity analysis, we analyze two approaches of FRAP data processing: (i) the integrated data approach commonly used by the FRAP community and (ii) the full (spatiotemporal) data case. The inertia of the common FRAP data processing methods almost prohibits promoting new approaches with few exceptions [10, 20, 21].

Our paper is organized as follows. After the introductory section we describe the problem and introduce the sensitivity analysis in Section 2. Then in Section 3, we develop a new theoretical approach allowing a data space reduction. Furthermore, in Section 4, we investigate a common way of treating the data (in the FRAP community), that is, using the so-called FRAP recovery curve in form of a time series (integrated data approach), and we state a proposition exhibiting larger sensitivities (and hence smaller confidence intervals) of the parameter identification problem with full data compared to that with integrated data. The novelty and

advantages of our approach as well as outlooks for further research are discussed in Section 5.

## 2. Preliminaries

*2.1. Reaction-Diffusion System.* Consider a model of reaction-diffusion system describing the fluorescent solute concentration in time and space. Right now we assume that two model parameters, (i) a diffusion coefficient  $D$  and (ii) a reaction factor  $R$ , are time-dependent and constant in space. The introduced reaction term reflects the continuous photobleaching during data acquisition and is proportional to the fluorescent particle concentration (which is proportional to the fluorescence signal  $u$  measured by CLSM); that is, it is described as a first-order reaction [10]. The governing equation for the spatiotemporal fluorescence signal  $u(x, t)$  is a diffusion equation with reaction term as follows:

$$\frac{\partial}{\partial t} u(x, t) = D\Delta u(x, t) - Ru(x, t),$$

$$x \in \Omega, t \in [0, T], \quad (1)$$

$$u(x, 0) = u_0(x), \quad x \in \Omega,$$

boundary conditions on  $\partial\Omega \times [0, T]$ .

Boundary conditions could be, for example,

$$u(x, t) = 0,$$

$$\text{or } \frac{\partial}{\partial n} u(x, t) = 0, \quad (2)$$

on  $\partial\Omega, t \in [0, T]$ .

We also consider the simplest case of unbounded domains,  $\Omega = \mathbb{R}^n$ , in which we set appropriate decay conditions at  $\|x\| \rightarrow \infty, t \in [0, T]$ . Equations (1) (and variants) are the basis for all the further analysis.

In the case of constant coefficients, the solution to this problem can be expressed by means of the Green function  $G(x, t; y)$  for the heat equation. Consider

$$\frac{\partial}{\partial t} G(x, t; y) = \Delta G(x, t; y), \quad x \in \Omega, t \in [0, T],$$

$$G(x, 0; y) = \delta(x - y), \quad x \in \Omega, \quad (3)$$

boundary conditions for  $G(x, t; y), \quad x \in \partial\Omega$ .

It is well-known that

$$\tilde{u}(x, t) = \int_{\Omega} G(x, t; y) u_0(y) dy \quad (4)$$

solves the problem for  $R = 0$  and  $D = 1$  and that

$$u(x, t) := e^{-Rt} \tilde{u}(x, Dt)$$

$$= e^{-Rt} \int_{\Omega} G(x, Dt; y) u_0(y) dy \quad (5)$$

solves (1) for constant parameters  $R, D$ .

Some frequently used cases are that of a diffusion in free space, that is, on the whole domain  $\mathbb{R}^n$  without boundary conditions. In that case, the Green function (in 1 and 2 spatial dimensions, resp.) is the well-known heat kernel. Consider

$$G(x, t; y) = \frac{1}{\sqrt{4\pi t}} e^{-(x-y)^2/4t}, \quad x, y \in \mathbb{R},$$

$$G(x, t; y) = \frac{1}{4\pi t} e^{-\|x-y\|^2/4t}, \quad x, y \in \mathbb{R}^2. \quad (6)$$

In FRAP experiments, the initial condition, that is, the first postbleach profile (with the background or prebleach signal subtracted), is often modeled as a Gaussian [4, 7, 8] (cf. Figure 1(b)) which leads in the one- or two-dimensional case to initial conditions of the form

$$u_0(x) = u_{0,0} e^{-2x^2/r_0^2} \quad (1D),$$

$$u_0(x) = u_{0,0} e^{-2\|x\|^2/r_0^2} \quad (2D), \quad (7)$$

where  $u_{0,0} < 0$  is the maximum depth at time  $t_0$  (i.e., for  $x = 0$ ) and  $r_0 > 0$  is the half-width of the bleach at height (depth)  $u_{0,0}e^{-2}$ ; see, for example, [7]. The explicit solution for  $u$  in the free space case is then given by

$$u(x, t) = e^{-Rt} u_{0,0} \frac{r_0}{\sqrt{r_0^2 + 8Dt}} e^{-2x^2/(r_0^2 + 8Dt)} \quad (1D),$$

$$u(x, t) = e^{-Rt} u_{0,0} \frac{r_0^2}{r_0^2 + 8Dt} e^{-2\|x\|^2/(r_0^2 + 8Dt)} \quad (2D). \quad (8)$$

*2.2. Parameter Identification Problem.* We now discuss the parameter identification problem, where we try to infer about the parameters  $D, R$  by using direct measurements of  $u$  in some space-time domain. That is, we assume that one of the following kind of data are observed:

$$u(x_i, t_i)_{i=1}^{N_{\text{data}}} \in \mathbb{R}^{N_{\text{data}}} \quad \text{Case Discrete Data (DD)}$$

$$u(x, t) \quad (9)$$

$$(x, t) \in B \subset \Omega \times [0, T] \quad \text{Case Continuous Data (CD)}.$$

The first formulation (DD) is a discrete one, and the second one (CD) is an idealized continuous version of it. In the second one, the set  $B$  is some observation (monitored) region, which does not have to be the full region  $\Omega \times [0, T]$ . Of course, if the data grid is small enough, the discrete data case can be seen as an approximation to the continuous case. We will use this at several instances, for example, for an approximation of sums in the discrete case by the corresponding integrals.

*Remark 1.* A FRAP data structure usually consists of a time sequence of rectangular matrices (2D fluorescence profiles, cf. Figure 1(a)), where each entry quantifies the fluorescence intensity at a particular point in a 2D domain (e.g., by a number between 0 and 255). Consider

$$u(x_j, t_j)_{j=1}^{N_x}, \quad j = 0 \cdots N_t, \quad (10)$$

where  $l$  is the spatial index uniquely identifying the position where the signal  $u$  is measured and  $j$  is the time index (the initial condition corresponds to  $j = 0$ ).

Moreover, the data points are uniformly distributed in both time (the time interval between two consecutive measurements is constant) and space (on an equidistant 1D or 2D mesh). Nevertheless, in the following we adopt the simplified notation consisting in using only one index  $i$  ( $i = 1 \cdots N_{\text{data}}$ ) for all data in the space-time domain.

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

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow \mathbb{R}^{N_{\text{data}}} \\ (D, R) &\longrightarrow u(x_i, t_i)_{i=1}^{N_{\text{data}}}, \end{aligned} \quad (11)$$

where  $u$  is as in (5). Our regression model is now

$$F(D, R) = \text{data}, \quad (12)$$

where the data are modeled as contaminated with additive white noise as follows:

$$\text{data} = F(D_T, R_T) + e = u(x_i, t_i)_{i=1}^{N_{\text{data}}} + (e_i)_{i=1}^{N_{\text{data}}}. \quad (13)$$

Here  $D_T, R_T$  denotes the true coefficients and  $e$  is a data error vector which we assume to be normally distributed with variance  $\sigma^2$ . Consider

$$(e_i)_{i=1}^{N_{\text{data}}} \in \mathbb{R}^{N_{\text{data}}}, \quad e_i = \mathcal{N}(0, \sigma^2). \quad (14)$$

Given some data, the aim of the parameter identification problem is to find  $D, R$  such that (12) is satisfied in some appropriate sense. Since (12) usually consists of an overdetermined system (there are more data points than unknowns), it cannot be expected that (12) holds with equality, but instead an appropriate notion of solution (which we adopt for the rest of the paper) is that of a least-squares solution  $D_c, R_c$  (with  $\|\cdot\|$  denoting the Euclidean norm on  $\mathbb{R}^{N_{\text{data}}}$ ). One has

$$\|F(D_c, R_c) - \text{data}\|^2 = \min_{D, R} \|F(D, R) - \text{data}\|^2. \quad (15)$$

The above defined parameter identification problem is usually ill-posed for nonconstant coefficients such that regularization has to be employed; see, for example, [13]. The solution of a practical example based on FRAP data was presented in [15]. If we *a priori* restrict the coefficients  $D, R$  to be constant, then the identification problem becomes well-posed.

Further, in Section 3, we solve and afterward discuss the problem of data space selection based on sensitivity analysis. The prerequisite for this study is the confidence intervals estimation which we introduce in the next subsection.

**2.3. Sensitivity Analysis and Confidence Intervals.** For the sensitivity analysis we require the Fréchet-derivative of the forward map  $F$ ; that is,

$$\begin{aligned} F' [D, R] &= \left( \frac{\partial}{\partial D} F(D, R) \quad \frac{\partial}{\partial R} F(D, R) \right) \\ &= \begin{pmatrix} \frac{\partial}{\partial D} u(x_1, t_1) & \frac{\partial}{\partial R} u(x_1, t_1) \\ \vdots & \vdots \\ \frac{\partial}{\partial D} u(x_{N_{\text{data}}}, t_{N_{\text{data}}}) & \frac{\partial}{\partial R} u(x_{N_{\text{data}}}, t_{N_{\text{data}}}) \end{pmatrix} \quad (16) \\ &\in \mathbb{R}^{N_{\text{data}} \times 2}. \end{aligned}$$

A corresponding quantity is the Fisher information matrix

$$M[D, R] = F' [D, R]^T F' [D, R] \in \mathbb{R}^{2 \times 2}. \quad (17)$$

According to Bates and Watts [22], we can estimate confidence intervals. Suppose we have computed  $(D_c, R_c)$  as least-squares solutions in the sense of (15). Let us define the residual as

$$\begin{aligned} \text{res}^2(D_c, R_c) &= \|F(D_c, R_c) - \text{data}\|^2 \\ &= \sum_{i=1}^{N_{\text{data}}} (\text{data}_i - u_{D_c, R_c}(x_i, t_i))^2. \end{aligned} \quad (18)$$

Then according to [22], it is possible to quantify the error between the computed parameters  $(D_c, R_c)$  and the true parameters  $(D_T, R_T)$ . In fact, we have an approximate  $1 - \alpha$  confidence interval as follows:

$$\begin{aligned} (D_c - D_T \quad R_c - R_T) M[D_c, R_c] \begin{pmatrix} D_c - D_T \\ R_c - R_T \end{pmatrix} \\ \leq 2 \frac{\text{res}^2(D_c, R_c)}{N_{\text{data}} - 2} f_{2, N_{\text{data}} - 2}(\alpha), \end{aligned} \quad (19)$$

where  $f_{m,n}(\alpha)$  corresponds to the upper  $\alpha$  quantile of the Fisher distribution with  $m$  and  $n$  degrees of freedom.

The parameter  $R$  can be either disregarded for its low value [4] or measured by another set of FRAP experiments. These new experiments allow the identification of parameters related to the fluorescence loss due to the bleaching during scanning; see, for example, [10]. Consequently, in case when  $R$  is either fixed (known) or disregarded ( $R = 0$ ), we only have  $D$  as unknown. Then a similar result holds and the Fisher information matrix collapses into the scalar quantity  $\sum_{i=1}^{N_{\text{data}}} ((\partial/\partial D)u(x_i, t_i))^2$ , and the confidence interval for full observations is described as follows:

$$\begin{aligned} (D_c - D_T)^2 \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ \leq \frac{\text{res}^2(D_c)}{N_{\text{data}} - 1} f_{1, N_{\text{data}} - 1}(\alpha), \end{aligned} \quad (20)$$

with

$$\text{res}^2(D_c) = \sum_{i=1}^{N_{\text{data}}} (\text{data}_i - u_{D_c}(x_i, t_i))^2. \quad (21)$$

*Remark 2.* In (20), several simplifications are possible. Note that according to our noise model, the residual term  $\text{res}^2(D_c)/(N_{\text{data}} - 1)$  is an estimator of the error variance [22] such that the approximation

$$\frac{\text{res}^2(D_c)}{N_{\text{data}} - 1} \sim \sigma^2 \quad (22)$$

holds for  $N_{\text{data}}$  being large. This approximation will be used in several instances in the following, and the term  $\text{res}^2(D_c)/(N_{\text{data}} - 1)$  in (20) can be viewed as rather independent of  $D_c$  or  $N_{\text{data}}$ .

Moreover, we remind the reader that the Fisher distribution with 1 and  $N_{\text{data}} - 1$  degrees of freedom converges to the  $\chi^2$ -distribution as  $N_{\text{data}} \rightarrow \infty$ . Hence, the term  $f_{1, N_{\text{data}}-1}(\alpha)$  can approximately be viewed as independent of  $N_{\text{data}}$  as well and of moderate size.

### 3. Theoretical Results on Data Space Selection

In this section we present the theoretical results concerning an (in further defined sense) optimal data space selection for the problem of parameter identification based on FRAP data. For all the results following, we always consider only the constant diffusivity  $D$  as unknown; set  $R = 0$ . Hence, (20) is the central estimate which we use in the sequel. We mention that most of the analysis can be extended to the case of two unknowns  $D, R$  as well. However, the case of an unknown nonconstant diffusivity is out of the scope of this paper.

*3.1. Data Reduction: Relevant versus Irrelevant Data.* The confidence interval estimate (20) gives us useful information on the quality of a least-squares estimate. A central ingredient in this estimate is the sensitivity (or the Fisher information matrix), which is the main factor controlling the error  $|D_c - D_T|$ . As a next step, starting from (20), we might change the data observation and consider the question if we can use less data without increasing the error too much.

Let us assume now that we have different kind of data (i.e., less data points, also called *reduced data* in the following). One has

$$u(\bar{x}_i, \bar{t}_i)_{i=1}^{\bar{N}_{\text{data}}} \in \mathbb{R}^{\bar{N}_{\text{data}}}, \quad (23)$$

where

$$\begin{aligned} \bar{N}_{\text{data}} &< N_{\text{data}}, \\ \{(\bar{x}_i, \bar{t}_i)\} &\subset \{(x_i, t_i)\}, \end{aligned} \quad (24)$$

and we compute a least-squares estimate  $\bar{D}_c$  using these data. Without loss of generality let us impose the following ordering of the data  $\{(\bar{x}_i, \bar{t}_i)_{i=1}^{\bar{N}_{\text{data}}}\} = \{(x_i, t_i)_{i=1}^{\bar{N}_{\text{data}}}\}$ , that is, that the new data are just the first  $\bar{N}_{\text{data}}$  values of the original

data. We generally impose the condition that  $\bar{N}_{\text{data}}$  is large enough such that certain estimators (like  $\text{res}^2(\bar{D}_c)/(\bar{N}_{\text{data}} - 1)$ ) are close to their expected value, that is, that a law of large numbers is reasonable. As a rule of thumb, we impose in all of the following that  $\bar{N}_{\text{data}} \geq 50$ . The case of extreme data reduction to a few points requires further more detailed analysis, which we omit here.

The corresponding confidence interval estimate (20) still holds with a (usually larger) confidence interval

$$\begin{aligned} &(\bar{D}_c - D_T)^2 \sum_{i=1}^{\bar{N}_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ &\leq \frac{\text{res}^2(\bar{D}_c)}{\bar{N}_{\text{data}} - 1} f_{1, \bar{N}_{\text{data}}-1}(\alpha). \end{aligned} \quad (25)$$

Note that for the residual we now have to take the Euclidean norm in  $\mathbb{R}^{\bar{N}_{\text{data}}}$  in the corresponding expression (21) (i.e., a sum of the form  $\sum_{i=1}^{\bar{N}_{\text{data}}}$ ).

Suppose we would like to have a similar confidence interval as before with the new data set  $(\bar{x}_i, \bar{t}_i)$  instead. We can thus estimate using (25)

$$\begin{aligned} &(\bar{D}_c - D_T)^2 \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \leq \frac{\text{res}^2(\bar{D}_c)}{\bar{N}_{\text{data}} - 1} \\ &\cdot f_{1, \bar{N}_{\text{data}}-1}(\alpha) + (\bar{D}_c - D_T)^2 \\ &\cdot \sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \leq \frac{\text{res}^2(\bar{D}_c)}{\bar{N}_{\text{data}} - 1} \\ &\cdot f_{1, N_{\text{data}}-1}(\alpha) \left[ \frac{f_{1, \bar{N}_{\text{data}}-1}(\alpha)}{f_{1, N_{\text{data}}-1}(\alpha)} \right] \\ &\cdot \left[ 1 + \frac{\sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} ((\partial/\partial D) u(x_i, t_i))^2}{\sum_{i=1}^{\bar{N}_{\text{data}}} ((\partial/\partial D) u(x_i, t_i))^2} \right]. \end{aligned} \quad (26)$$

Let us further introduce an upper bound  $\eta > 0$  such that

$$\eta \geq \frac{\sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} ((\partial/\partial D) u(x_i, t_i))^2}{\sum_{i=1}^{\bar{N}_{\text{data}}} ((\partial/\partial D) u(x_i, t_i))^2}, \quad (27)$$

then we finally find a confidence interval using the reduced data. It will be announced in the following lemma.

**Lemma 3.** *Let  $\bar{N}_{\text{data}}$  be the number of reduced data points and suppose that  $\eta$  is chosen according to (27). Then using reduced data, one finds a confidence interval of the form*

$$\begin{aligned} &(\bar{D}_c - D_T)^2 \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \leq \left[ \frac{f_{1, \bar{N}_{\text{data}}-1}(\alpha)}{f_{1, N_{\text{data}}-1}(\alpha)} \right] \\ &\cdot (1 + \eta) \left( \frac{\text{res}^2(\bar{D}_c, R_c)}{\bar{N}_{\text{data}} - 1} f_{1, N_{\text{data}}-1}(\alpha) \right). \end{aligned} \quad (28)$$

*Remark 4.* The relevance of this observation is that we have almost the same type of estimates as in the full data case (20), but we might ignore irrelevant data points. The Fisher distribution converges to the  $\chi^2$ -distribution as  $N_{\text{data}} \rightarrow \infty$  and thus it holds that  $f_{1, \bar{N}_{\text{data}}-1}(\alpha)/f_{1, N_{\text{data}}-1}(\alpha) \sim 1$  if both  $N_{\text{data}}$  and  $\bar{N}_{\text{data}} \rightarrow \infty$ . For reasonable values of  $\alpha$  and  $\bar{N}_{\text{data}}$  (e.g.,  $\alpha = 0.05$  and  $\bar{N}_{\text{data}} \geq 50$ ), it can be verified that  $f_{1, \bar{N}_{\text{data}}-1}(\alpha)/f_{1, N_{\text{data}}-1}(\alpha) \leq 1.05$ . This situation corresponds to our initial assumption of sufficiently large  $\bar{N}_{\text{data}}$ . Note that a similar analysis as below could be done if this assumption is not valid by incorporating the factor  $f_{1, \bar{N}_{\text{data}}-1}(\alpha)/f_{1, N_{\text{data}}-1}(\alpha)$  into the analysis.

Thus, in practice and for the later analysis we replace the estimate (28) by the simpler one (forgetting the first factor). Consider

$$\begin{aligned} & (\bar{D}_c - D_T)^2 \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & \leq (1 + \eta) \left( \frac{\text{res}^2(\bar{D}_c, R_c)}{\bar{N}_{\text{data}} - 1} f_{1, N_{\text{data}}-1}(\alpha) \right). \end{aligned} \quad (29)$$

*Remark 5.* The previous Lemma 3 can be seen as a criterion for selecting  $\bar{N}_{\text{data}}$ . If this  $\eta$  is chosen as an upper bound, then (27) is relating  $\bar{N}_{\text{data}}$  to  $\eta$ . Then this will be the main condition for choosing  $\bar{N}_{\text{data}}$  (as small as possible such that (27) holds). With this choice and if  $\eta$  is small, we basically obtain almost the same confidence interval than when using the full data. Hence, out of the original data  $u(x_i, t_i)_{i=1}^{N_{\text{data}}}$  we can pick a subset  $u(\bar{x}_i, \bar{t}_i)_{i=1}^{\bar{N}_{\text{data}}}$  which contains almost the same information. If (27) holds with small  $\eta$ , we thus call the new data  $u(\bar{x}_i, \bar{t}_i)_{i=1}^{\bar{N}_{\text{data}}}$  the relevant data and the complement  $\{u(x_i, t_i)_{i=1}^{N_{\text{data}}}\} \setminus \{u(\bar{x}_i, \bar{t}_i)_{i=1}^{\bar{N}_{\text{data}}}\}$  the irrelevant data.

Let us mention, following Remark 2, that the expressions  $\text{res}^2(\bar{D}_c)/(\bar{N}_{\text{data}} - 1)$  and  $\text{res}^2(D_c)/(N_{\text{data}} - 1)$  are both similar to  $\sigma^2$  for  $\bar{N}_{\text{data}}$  and  $N_{\text{data}}$  are large. More precisely, the estimator  $\text{res}^2(D_c)/(N_{\text{data}} - 1)$  has variance  $\sigma^4/N_{\text{data}}$ , and hence the ratio of these two terms is roughly of the order of  $1 + \sigma^4/\bar{N}_{\text{data}}$ . With our assumption of large data sizes,  $\bar{N}_{\text{data}} \geq 50$ , and for reasonable  $\sigma$ , the two expressions can hence be considered equal in our analysis. This explains, why we considered the right-hand sides of the corresponding estimates (20) and (28) (resp., (29)) almost equal. In fact, our analysis is rather independent of the  $\text{res}^2(\bar{D}_c)/(\bar{N}_{\text{data}} - 1)$  and  $\sigma$  (as long as we impose the condition of  $\bar{N}_{\text{data}}$  being large as above) because we compare two situations (full data and reduced data) with similar  $\sigma$  and the corresponding factors cancel out. The dependence of the confidence intervals itself on the noise variance is, of course, well known and (in the well-posed case) linear; see, for example, [6, 9] and Figure 4.

**3.2. Data Reduction Based on the Data Noise Level.** There is also the possibility of using other selection criteria as well for the data reduction step. The following one is based on comparing the sensitivity with the noise level. In general, for

deterministic errors, the data noise level refers to the norm of the noise. In our case, that is, that of stochastic errors, we identify the noise level with the noise variance  $\sigma^2$ .

**Lemma 6.** Let  $\bar{N}_{\text{data}}$  be the number of reduced data points and let  $|(D_c - D_T)/D_c| \leq 1$  hold. (This means that the relative difference  $|D_c - D_T|$  with respect to  $D_c$  is less than 100%.) Furthermore, we assume the following selection criterion with some freely selected parameter  $\tau > 0$ :

$$D_c^2 \sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \leq \tau \sigma^2. \quad (30)$$

Then using reduced data, we find a confidence interval of the form

$$\begin{aligned} & (\bar{D}_c - D_T)^2 \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & \leq \left( \frac{\text{res}^2(\bar{D}_c, R_c)}{\bar{N}_{\text{data}} - 1} f_{1, \bar{N}_{\text{data}}-1}(\alpha) + \tau \sigma^2 \right). \end{aligned} \quad (31)$$

*Proof.* The proof follows the steps of that preceding Lemma 3. We obtain the estimate in Lemma 6 by using assumptions (30) and  $|(D_c - D_T)/D_c| \leq 1$  in the second part of (26) as

$$\begin{aligned} & (\bar{D}_c - D_T)^2 \sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & = \left( \frac{\bar{D}_c - D_T}{D_c} \right)^2 D_c^2 \sum_{i=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & \leq \tau \sigma^2, \end{aligned} \quad (32)$$

which immediately yields the result.  $\square$

*Remark 7.* If the approximation of residual (22) is valid and if the ratio of  $f_{1, \bar{N}_{\text{data}}-1}(\alpha)/f_{1, N_{\text{data}}-1}(\alpha) \sim 1$  as in Lemma 3, then Lemma 6 leads to a confidence interval of the order

$$\begin{aligned} & (\bar{D}_c - D_T)^2 \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & \sim \sigma^2 (f_{1, \bar{N}_{\text{data}}-1}(\alpha) + \tau) \\ & \sim \left( 1 + \frac{\tau}{f_{1, \bar{N}_{\text{data}}-1}(\alpha)} \right) f_{1, N_{\text{data}}-1}(\alpha) \frac{\text{res}^2(D_c, R_c)}{N_{\text{data}} - 1}, \end{aligned} \quad (33)$$

which is again of the same order as the full data case; the length of the confidence interval in the reduced data case is multiplied by the factor  $(1 + \tau/f_{1, \bar{N}_{\text{data}}-1}(\alpha))$  compared to the full data case.

*Remark 8.* The previous analysis can be extended to the case of an unknown reaction rate  $R$  as well. The selection criterion corresponding to Lemmas 3 and 6 should be formulated as

$$\eta \geq \frac{\|M[D_c, R_c] - M[\bar{D}_c, \bar{R}_c]\|}{\|M[\bar{D}_c, \bar{R}_c]\|}, \quad (34)$$

$$\|M[D_c, R_c] - M[\bar{D}_c, \bar{R}_c]\| \leq \tau\sigma^2, \quad (35)$$

respectively.

*Example 9.* In this example we underline the previous theoretical results by numerical computations. To do so, we compute a least-squares estimate  $D_c$  using different kind of data. In a first experiment we consider a rectangular spatiotemporal data grid with space interval  $x_i \in [-4, 4]$  and time interval  $t_i \in [0, 6]$ . The grid size in both space and time direction was set to  $\Delta x = \Delta t = 0.05$ . We simulated data by assuming an exact diffusion coefficient  $D_T = 2$  with bleach radius  $r_0 = 0.1$  and computed the data for the 1D case by (8). These data were perturbed by normally distributed additive noise (white Gaussian noise) with standard deviation  $\sigma = 0.05$ . Based on these data we computed a least-squares estimate  $D_c$  of the diffusion coefficient using a Gauss-Newton procedure. We refer to this setup as the full data case.

Next, we computed a similar estimate using less data. For this, we considered the data point selection criterion (27) for various values of  $\eta = 1, 0.3, 0.1$ . The data regions were selected by first ordering all data points according to their sensitivities  $((\partial/\partial D)u(x_1, t_1))^2 \leq ((\partial/\partial D)u(x_2, t_2))^2 \leq \dots$ . Then we selected the reduced data from the region of points  $(x_i, t_i), i, 1, \dots, \bar{N}_{\text{data}}$  using this ordering and the smallest  $\bar{N}_{\text{data}}$  such that (27) holds for the choice of  $\eta$ . In a sense, we take the most sensitive data points as relevant data. In Figure 2(a), the gray region corresponds to that region where the data points were selected and the complementary region, the white one, represents irrelevant data that can be neglected without much loss of quality. A similar process (using the same sensitivity ordering) was performed using selection criterion (30) with values of  $\tau = 20, 5, 1$ . The corresponding data regions are displayed in Figure 2(b). Below of each relevant data region, the ratio of  $\bar{N}_{\text{data}}/N_{\text{data}}$  is shown; that is,  $(1 - \bar{N}_{\text{data}}/N_{\text{data}}) \cdot 100\%$  represents the percentage of data reduction. For instance, in the first row, rightmost plot,  $\eta = 0.1$ , only 50% of the original data are used.

Having identified several regions of relevant data using different selection criteria, we considered the quality of the computed estimate  $D_c$  using the reduced data. For the seven cases (full data and six cases of reduced data), we sampled 1000 times a normally distributed noise vector  $e$  (with fixed standard deviation  $\sigma = 0.05$ ), and for each of the noise realizations, we computed a least-squares fit by a Gauss-Newton iteration and computed the squared error  $|D_c - D_T|^2$ . This yields 1000 samples out of the error distribution of  $|D_c - D_T|^2$ . For the seven cases, a boxplot of this error can be seen in Figure 2(c). Here the box interior line corresponds to the mean of the  $|D_c - D_T|^2$  (over 1000 realizations) and

the edges of the box in each column represents the 25th and 75th percentiles. The cross marks  $+$  are considered outliers. The size of the box (more precisely the top of the box) can be interpreted as an indicator of the size of a confidence interval.

This plot corresponds quite well with our theoretical findings. For instance, the case  $\eta = 1$ , in the second column of the boxplot, corresponds to the case of using only 10% of the original data, that is, 90% of data reduction. According to Lemma 3, the theoretical confidence interval bound is roughly  $1 + \eta = 2$  times larger compared to the full data case. And indeed, we observe the box in the second column in Figure 2(c) being roughly 2 times larger compared to the full data case in column one. The other cases with  $\eta = 0.3$  and  $\eta = 0.1$  behave similarly. Although the boxes in the last three columns for  $\tau = 20, 5, 1$  are within the theoretical bound in Lemma 6, it seems that this bound is not very precise, since the actual error is quite smaller than the bound in (31). In that respect, the data selection method of Lemma 3 is more accurate and preferable.

Note that for the case  $\eta = 0.1$  (fourth column), we have virtually the same error distribution as for the full data case even though we have only used 50% of the original data.

There is a rather esthetic issue of the lotus flower-like plots in Figure 2(a); what is the meaning of these lotus-like grey regions? The answer is simple: those regions exhibit the sensitivity to the fluorescence loss in photobleaching (FLIP); see [18] for review and [23] for a more creative exploration of the FRAP and FLIP synergic effect.

*Remark 10.* In Example 9, we have—in a sense—selected the optimal region by ordering the sensitivities. However, the data selection criteria of Lemmas 3 and 6 do not depend on this ordering but can be performed in different situations. In fact, in practice, one probably wants to select a rectangular region as a relevant data region instead of the complicated one in Figure 2. This is of course possible in Lemmas 3 and 6. One can compare all subregions of rectangular form and accept those which satisfy the criterion in the above lemmas. Such a concept will be pursued in the following subsections.

*Remark 11.* Another way of reducing the amount of data could be to use a coarser data grid; that is, instead of  $u(x_i, t_i)$ , one might perform a subsampling and just use every second grid point (in space or time direction). However, such an idea does not fit into our context because a subsampling will in general increase the confidence interval significantly. Looking at (25), one can observe that a subsampling by taking every second data point will lead to Fréchet-derivative which is about half the size of the original one and thus leading to a doubling of the confidence intervals. On the contrary, in our framework we consider cases where the confidence intervals stay almost the same. Thus subsampling will lead to an  $\eta$  which is not of the order of 1. Still, such a procedure makes sense, if one has an idea of the wanted size of the confidence intervals. Then the grid size can be selected such that estimate (25) fits to one's needs. For a rectangular grid on say a space-time region  $[-L/2, L/2] \times [0, T]$ , this requires an estimate of

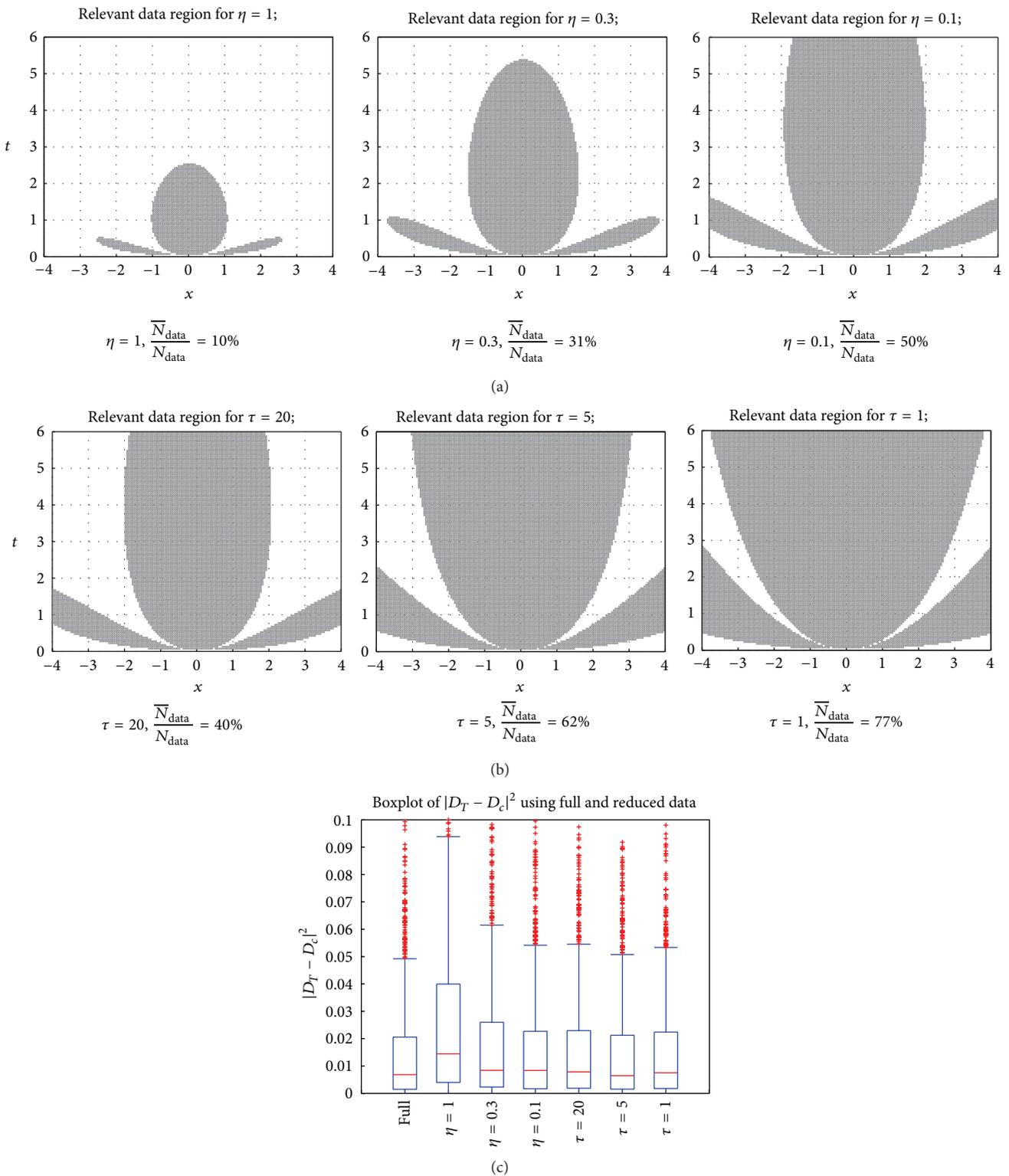


FIGURE 2: Examples of data reduction. Gray color indicates the data space, where the relevant data are taken. White color represents the data regions of irrelevant data points. (a) Relevant data region using (27) with  $\eta = 1, 0.3, 0.1$ . (b) Relevant data region using (30) with  $\tau = 20, 5, 1$ . (c) Boxplot of the distribution of the squared error  $|D_T - D_c|^2$  based on 1000 noisy signal samples using full and reduced data; cross marks (+) indicates the outliers. The data were simulated using (8), with an exact diffusion coefficient  $D_T = 2$ , bleach radius  $r_0 = 0.1$ , and additive white Gaussian noise with standard deviation  $\sigma = 0.05$ .

$\sum_{i=1}^{N_{\text{data}}} ((\partial/\partial D)u(x_i, t_i))^2$ , which can roughly be obtained by the asymptotics

$$\begin{aligned} & \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & \sim \frac{1}{\Delta x \Delta t} \int_{[-L/2, L/2] \times [0, T]} \left( \frac{\partial}{\partial D} u(x, t) \right)^2 dx dt \quad (36) \\ & = \mathcal{O} \left( \frac{1}{\Delta x \Delta t} \right) \end{aligned}$$

as  $\Delta x, \Delta t \rightarrow 0$ . Thus from estimates of the integral on the right-hand side,  $\Delta x, \Delta t$  can be adapted to reach a desired confidence.

**3.3. Data Reduction for the Free Space Case.** We now study the data reduction procedure in view of practical relevance. That is, we want to find simple formulas for selecting a rectangular spatiotemporal region which yield approximately similar error bounds than when using all the spatial points  $x \in \mathbb{R}^n$  as data.

Let us consider the problem on free space with  $R = 0$  and unknown  $D$  and a Gaussian as in (7) as initial condition. From the solutions (8) in the one- and two-dimensional case, we obtain by some calculations for the derivative

$$\begin{aligned} \frac{\partial}{\partial D} u(x, t) &= -u_{0,0} \frac{r_0}{(r_0^2 + 8Dt)^{5/2}} \\ &\cdot e^{-2x^2/(r_0^2 + 8Dt)} 4t (r_0^2 + 8Dt - 4x^2), \quad (1D) \end{aligned} \quad (37)$$

$$\begin{aligned} \frac{\partial}{\partial D} u(x, t) &= -u_{0,0} \frac{r_0^2}{(r_0^2 + 8Dt)^3} \\ &\cdot e^{-2\|x\|^2/(r_0^2 + 8Dt)} 8t (r_0^2 + 8Dt - 2\|x\|^2) \quad (2D). \end{aligned}$$

By an appropriate scaling based on the experimental design factor  $r_0$  (yielding the new spatial dimensionless variable  $z := x/r_0$ ) and on the diffusion coefficient  $D$  (yielding the second dimensionless variable  $s := tD/r_0^2$ ) we may simplify the expressions above.

**Lemma 12.** *In the setting as before one can express the sensitivity as*

$$\begin{aligned} \left| \frac{\partial}{\partial D} u(x, t) \right|^2 &= \frac{|u_{0,0}|^2}{|D|^2} H_1 \left( \frac{x}{r_0}, \frac{tD}{r_0^2} \right) \quad (1D), \\ \left| \frac{\partial}{\partial D} u(x, t) \right|^2 &= \frac{|u_{0,0}|^2}{|D|^2} H_2 \left( \frac{x}{r_0}, \frac{tD}{r_0^2} \right) \quad (2D) \end{aligned} \quad (38)$$

with the functions

$$\begin{aligned} H_1(z, s) &= \frac{16s^2}{(1+8s)^3} e^{-4z^2/(1+8s)} \left( 1 - \frac{4z^2}{1+8s} \right)^2, \\ H_2(z, s) &= \frac{64s^2}{(1+8s)^4} e^{-4\|z\|^2/(1+8s)} \left( 1 - \frac{1}{2} \frac{4\|z\|^2}{1+8s} \right)^2. \end{aligned} \quad (39)$$

In the one-dimensional case let us further assume the usual situation that full data are given on a space-time cylinder  $Q = [-L/2, L/2] \times [0, T]$ , where  $[-L/2, L/2]$  is the space interval of observations and  $[0, T]$  is the time interval as follows:

$$\begin{aligned} (x_i, t_i) &\in \left( \left[ -\frac{L}{2}, \frac{L}{2} \right] \times [0, T] \right) \\ &\cap \{ (k\Delta x, l\Delta t) \mid (k, l) \in N \}. \end{aligned} \quad (40)$$

Here  $\Delta x$  and  $\Delta t$  are the spacings of the grid in space and time direction, respectively. The sensitivity can be approximated by an integral

$$\begin{aligned} & \sum_{i=1}^{N_{\text{data}}} \left( \frac{\partial}{\partial D} u(x_i, t_i) \right)^2 \\ & \sim \frac{1}{\Delta x \Delta t} \int_{-L/2}^{L/2} \int_0^T \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt. \end{aligned} \quad (41)$$

Forgetting the factor  $1/\Delta x \Delta t$  (which cancels out anyway later), we find

$$\begin{aligned} & \int_{-L/2}^{L/2} \int_0^T \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt \\ &= \frac{|u_{0,0}|^2}{|D_c|^2} \int_{-L/2}^{L/2} \int_0^T H_1 \left( \frac{x}{r_0}, \frac{tD_c}{r_0^2} \right) dx dt \\ &= \frac{|u_{0,0}|^2 r_0^3}{|D_c|^3} \int_0^{TD_c/r_0^2} \int_{-L/2r_0}^{L/2r_0} H_1(z, s) dz ds \\ &= \frac{|u_{0,0}|^2 r_0^3}{|D_c|^3} K_1 \left( \frac{L}{2r_0}, \frac{TD_c}{r_0^2} \right) \end{aligned} \quad (42)$$

with

$$\begin{aligned} K_1(y, t) &= \int_0^t \int_{-y}^y H_1(z, s) dz ds \\ &= 2 \int_0^t \int_0^y H_1(z, s) dz ds \\ &= 2 \int_0^t 4s^2 y \frac{1+8s-8y^2}{(1+8s)^4} e^{-4(y^2/(1+8s))} \\ &\quad + s^2 \frac{3\sqrt{\pi}}{(1+8s)^{5/2}} \operatorname{erf} \frac{2y}{\sqrt{1+8s}} ds, \end{aligned} \quad (43)$$

and the error function  $\operatorname{erf}(x) = (2/\sqrt{\pi}) \int_0^x e^{-t^2} dt$ .

Analogously in the two-dimensional case, we can consider data on a space-time cylinder  $Q = \{\|x\| \leq r\} \times [0, T]$ , yielding that

$$\begin{aligned} & \int_{\|x\| \leq r} \int_0^T \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt \\ &= \frac{|u_{0,0}|^2}{|D_c|^2} \int_{\|x\| \leq r} \int_0^T H_2 \left( \frac{x}{r_0}, \frac{tD_c}{r_0^2} \right) dx dt \quad (44) \\ &= 2\pi \frac{|u_{0,0}|^2 r_0^4}{|D_c|^3} K_2 \left( \frac{r}{r_0}, \frac{TD_c}{r_0^2} \right) \end{aligned}$$

with

$$\begin{aligned} K_2(y, t) &= \int_0^t \int_0^y H_2(z, s) z dz ds = \int_0^t \frac{4s^2}{(1+8s)^3} \left( 1 \right. \\ &\quad \left. + e^{-4y^2/(1+8s)} \left( -1 + 4 \frac{y^2}{1+8s} - 8 \frac{y^4}{(1+8s)^2} \right) \right) ds. \quad (45) \end{aligned}$$

A particular case happens if  $L \rightarrow \infty$  in the above integrals. Then

$$\begin{aligned} & \int_{-\infty}^{\infty} \int_0^T \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt \\ &= K_1 \left( \infty, \frac{TD_c}{r_0^2} \right) = 2 \int_0^{TD_c/r_0^2} \frac{3\sqrt{\pi}}{(1+8s)^{5/2}} ds \quad (46) \\ &= \frac{\sqrt{\pi} \left( 1 + 12 (TD_c/r_0^2) + 24 (TD_c/r_0^2)^2 \right)}{16 (1 + 8 (TD_c/r_0^2))^{3/2}} - \frac{\sqrt{\pi}}{16}, \end{aligned}$$

respectively,

$$\begin{aligned} & \int_{\mathbb{R}^2} \int_0^T \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt \quad (47) \\ &= K_2 \left( \infty, \frac{TD_c}{r_0^2} \right) = \int_0^{TD_c/r_0^2} \frac{4s^2}{(1+8s)^3} ds. \end{aligned}$$

The data reduction process involves estimates of the sensitivities and hence of  $K_1(y, t), K_2(x, t)$  as in (27). The previous calculations lead to the following lemma, which allows for data reduction in the spatial variable compared to full spatial information; see Example 14 below.

**Lemma 13.** *Let  $K_1(y, t)$  and  $K_2(y, t)$  be defined as in (43) and (45), respectively. Let  $t > 0$  and  $\beta > 1$ . If*

$$y \geq \beta \sqrt{8t + 1}, \quad (48)$$

then

$$\begin{aligned} \frac{K_1(y, t)}{K_1(\infty, t)} &\geq f_1(\beta), \\ \frac{K_2(y, t)}{K_2(\infty, t)} &\geq f_2(\beta), \quad (49) \end{aligned}$$

where

$$\begin{aligned} f_1(\beta) &= \operatorname{erf}(2\beta) + \frac{4}{3\sqrt{\pi}} e^{-4\beta^2} \beta (1 - 8\beta^2), \\ f_2(\beta) &= 1 + e^{-4\beta^2} (-1 + 4\beta^2 - 8\beta^4). \quad (50) \end{aligned}$$

*Proof.* With (48) we have that  $y \geq \beta \sqrt{8s + 1}$  for all  $s \leq t$ . With  $\xi = y/\sqrt{1 + 8s} \geq \beta$ , (43) reads

$$\begin{aligned} K_1(y, t) &= 2 \int_0^t s^2 \frac{3\sqrt{\pi}}{(1+8s)^{5/2}} \\ &\quad \cdot \left( \operatorname{erf}(2\xi) + \frac{4}{3\sqrt{\pi}} e^{-4\xi^2} \xi (1 - 8\xi^2) \right) ds \geq f_1(\beta) \quad (51) \\ &\quad \cdot 2 \int_0^t s^2 \frac{3\sqrt{\pi}}{(1+8s)^{5/2}} = f_1(\beta) K_1(\infty, t), \end{aligned}$$

where we used the fact that  $f$  is monotonically increasing for  $\beta \geq 1$ . (This can easily be seen from its derivative.) The same proof holds with  $f_2$  in place of  $f_1$ . Note that  $f_2$  is again monotonically increasing for  $\beta \geq 1$ .  $\square$

This lemma leads to simple rules for selecting the length of data intervals with respect to the spatial dimension, as the following example explains.

*Example 14.* Let us now again exemplify the data reduction procedure. Suppose we want to find the relevant data on a space-time cylinder  $Q = [-L/2, L/2] \times [0, T]$  which yields a comparable confidence interval to full data given on the whole spatial domain  $[-L'/2, L'/2] \times [0, T]$  with  $L'$  being large and with equal grid spacings  $\Delta t, \Delta x$ . Assuming  $L'$  being large, we can approximately view the full data interval as  $[-L'/2, L'/2] \sim \mathbb{R}$  and in the following treat this as the case when the full data are given on the whole space interval.

The condition we impose now is (27), with some chosen small  $\eta$ . We have the situation that the ideal  $N_{\text{data}}$  corresponds to data on  $\mathbb{R} \times [0, T]$ , while the actual data are given on  $[-L/2, L/2] \times [0, T]$ . Condition (27), assuming that the data points are dense such that the sums can be approximated by integrals, can be stated as

$$\begin{aligned} \eta & \geq \frac{\int_0^T \int_{-\infty}^{\infty} \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt - \int_0^T \int_{-L/2}^{L/2} \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt}{\int_0^T \int_{-L/2}^{L/2} \left| \frac{\partial}{\partial D} u(x, t) \right|^2 dx dt}. \quad (52) \end{aligned}$$

Rearranging terms and with the notation as before we come to the condition

$$\eta \geq \frac{K_1(\infty, TD_c/r_0^2) - K_1(L/2r_0, TD_c/r_0^2)}{K_1(L/2r_0, TD_c/r_0^2)}, \quad (53)$$

or equivalently

$$\frac{K_1(L/2r_0, TD_c/r_0^2)}{K_1(\infty, TD_c/r_0^2)} \geq \frac{1}{1 + \eta}. \quad (54)$$

We are now in the situation, where we can apply Lemma 13. Using (48) with  $y = L/2r_0$  and  $t = TD_c/r_0^2$  and given  $\eta$ , we have to find a  $\beta > 1$  such that  $f_1(\beta) \geq 1/(1 + \eta)$ , respectively,

$$\frac{1}{f_1(\beta)} - 1 \leq \eta. \tag{55}$$

This inequality can be solved by numerical means; for instance, if we allow for a 1% increase of the confidence interval, we take  $\eta = 0.01$  and find numerically a value of  $\beta = 1.35$ . For  $\eta = 0.001$  (0.1% increase) we find a value  $\beta = 1.59$ . Thus with the above setting of variables, we have the following recommendation for the length of data interval  $L$ :

$$\begin{aligned} \frac{L}{2} &\geq 3.82\sqrt{TD_c + \frac{1}{8}r_0^2} \quad \text{for } \eta \leq 0.01, \\ \frac{L}{2} &\geq 4.5\sqrt{TD_c + \frac{1}{8}r_0^2} \quad \text{for } \eta \leq 0.001, \end{aligned} \tag{56}$$

which leads to the confidence intervals that are almost identical to the case of full spacial data.

A similar procedure holds in the two-dimensional case. If instead of full data on  $\mathbb{R}^2 \times [0, T]$  we want to take only relevant data on a cylinder  $Q = \{\|x\| \leq r\} \times [0, T]$ , we have to find a  $\beta$  such that (55) holds with  $f_2$  instead of  $f_1$ . A numerical examination of this function reveals the recommendation for the spatial radius of the data-cylinder. One has

$$\begin{aligned} r &\geq 9.6\sqrt{TD_c + \frac{1}{8}r_0^2} \quad \text{for } \eta \leq 0.01, \\ r &\geq 10.7\sqrt{TD_c + \frac{1}{8}r_0^2} \quad \text{for } \eta \leq 0.001. \end{aligned} \tag{57}$$

Note that in these examples we only restrict the spatial coordinates while the time interval  $[0, T]$  for the full data case and the restricted data case is the same. The question comes up if it is possible to reduce the time interval as well and to find a smaller interval  $[0, \bar{T}]$  such that (27) holds with a small  $\eta$  and hence yielding similar confidence intervals. Unfortunately it is not possible to do so without some loss in precision of the resulting parameter estimates. In fact, suppose that the spatial interval  $[-L/2, L/2]$  is chosen as in (56), then we can approximate the sensitivity on the interval  $[-L/2, L/2] \times [0, T]$  well by  $K_1(\infty, T)$ . Hence (27) holds using this approximation if

$$\eta \geq \frac{K_1(\infty, T) - K_1(\infty, \bar{T})}{K_1(\infty, \bar{T})}. \tag{58}$$

According to Lemma 3, the ratio of the confidence interval using data on  $[-L/2, L/2] \times [0, T]$  and  $[-L/2, L/2] \times [0, \bar{T}]$  is then approximately given by

$$\begin{aligned} &\frac{|\bar{D}_c - D_T|_{\text{data on } [-L/2, L/2] \times [0, T]}^2}{|\bar{D}_c - D_T|_{\text{data on } [-L/2, L/2] \times [0, \bar{T}]}^2} \\ &\sim (1 + \eta) \sim \frac{K_1(\infty, T)}{K_1(\infty, \bar{T})}. \end{aligned} \tag{59}$$

The last ratio can be easily evaluated using (46). It turns out that this ratio is of the order of 1 only if  $T \sim \bar{T}$ . Thus it is always better to have a larger time interval. Still, if one can quantify the ‘‘costs’’ of using a large time interval and adds these to the ratio in (59), this can lead to an optimization problem in which case an optimal time interval  $[0, \bar{T}]$  of data measurements can be easily calculated.

*Example 15.* Let us illustrate the findings of this subsection by numerical examples. The main point in the previous analysis was that a very large spatial data domain  $x \in \mathbb{R}$  can be replaced without loss of quality by a smaller one  $[-L/2, L/2]$  using formula (56). For the results in Figure 3, we first considered data given on a rectangular grid with large spatial size. The full data case corresponds to data given on the spatial interval  $x_i \in [-25, 25]$  and on the time interval  $t_i \in [0, 10]$ . The spacing of the grid was  $\Delta x = 0.05$  in space direction and  $\Delta t = 0.05$  in time direction. For all the computations in this example we used an exact diffusivity  $D_T = 2$  and normally distributed additive random noise with  $\sigma = 0.05$ . As in Example 9, we computed an estimate  $D_c$  for the diffusivity by using different kind of data. For each data region we used 1000 samples of noise yielding the same number of squared errors  $|D_c - D_T|^2$ .

We considered seven different kinds of rectangular data regions. The corresponding data regions (with labels (I)–(VII)) are indicated in Figure 3(a) by different shadings. The corresponding boxplot of the error is given in Figure 3(b).

The first boxplot in Figure 3(b) (labeled as (I)) corresponds to the full data case with data given on the full grid on  $(x_i, t_i) \in [-25, 25] \times [0, 10]$ . For the second boxplot (labeled as (II)) we used reduced data on an interval  $[-L/2, L/2] \times [0, T]$  with  $L$  computed by (56) which yields  $L/2 \sim 13.5$ . Note that the time interval here is the same as that for the full data case. According to the above analysis such a procedure with (56) should not change the confidence interval significantly. Indeed, we observe in Figure 3(b) that the cases (I) and (II) are almost equal. Note that in case (II), we only use 54% of the full data with similar results, which again indicates the big amount of irrelevant data in the full data case.

The plots on the right, labeled as (III)–(VII), correspond to a choice of a smaller time interval with the corresponding length of the spatial interval  $L$  given by (56). We have already indicated that it is always better to have a larger time interval, because a smaller time interval will yield a larger confidence interval according to (59).

In case (III) we considered a time interval with  $T = 0.335(r_0^2/D) \sim 0.04$  which yields a very small data region indicated by the (hardly visible) darkest color in the right plot close to the origin of space-time coordinate system. Case (IV) corresponds to  $T = 2(r_0^2/D) \sim 0.25$ , case (V) corresponds to  $T = 20(r_0^2/D) \sim 2.5$ , and cases (VI) and (VII) correspond to  $T = 5$  and  $T = 7.5$ , respectively. The according spatial interval is set by (56). The boxplots on the right-hand side correspond to the theoretical findings above. Note that case (III) has a too large error to fit to the scaling of the plot (hence, there is no corresponding box in Figure 3(b)). We observe that all the boxes in cases (III)–(VII)

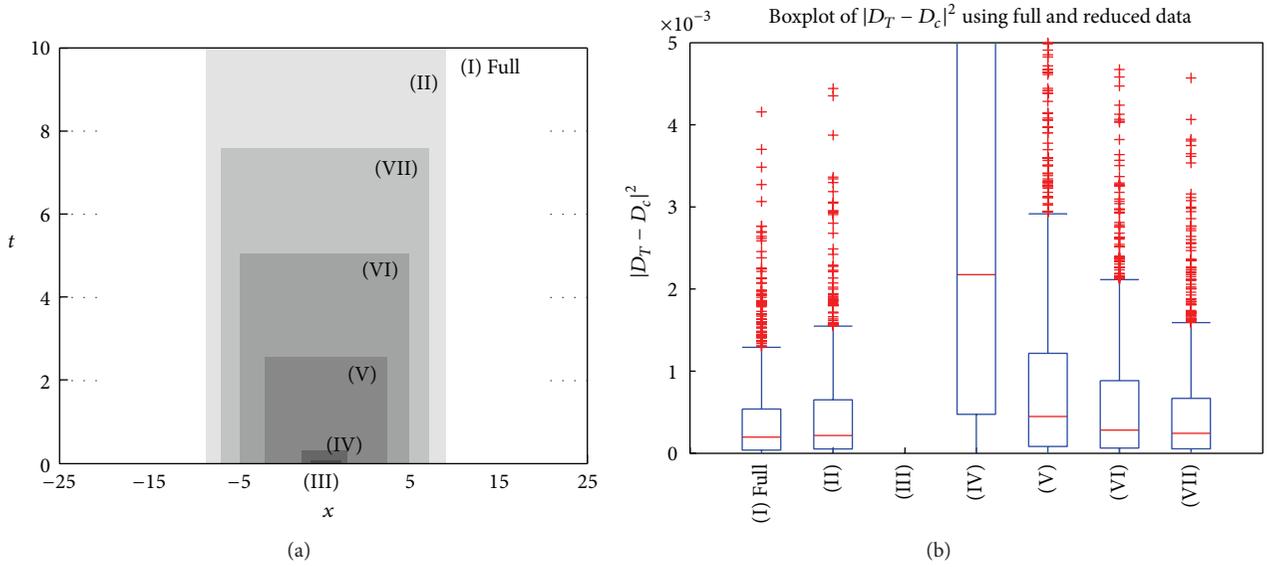


FIGURE 3: Comparison of the squared error  $|D_T - D_c|^2$  for 7 rectangular spatiotemporal regions. (a) Data regions for the cases (I)–(VII); full data are represented by the white region  $[-25, 25] \times [0, 10]$ , case (I). Reduced data, in cases (II)–(VII), are indicated by different shadings. (b) Boxplot of the error  $|D_T - D_c|^2$  corresponding to the use of data from the respective regions.

have larger confidence intervals than those in the full data case (I). Only the last case (VII) is similar to the one in the full data case (I). This shows that a reduction of the time interval (Another issue to be considered is the enlargement of the time interval between two consecutive measurements as the FRAP experiment proceeds.) has to be paid by an enlargement of the confidence intervals.

In Figure 4 we performed the same experiments as in Figure 3 for some selected regions, but with different noise variances  $\sigma^2$ . The first 5 plots correspond to the confidence intervals of Figure 3(b) for regions (I), (II), (IV), (V), and (VII), but now with a different  $\sigma$  for each plot. From left to right and top to bottom we have  $\sigma = (0.01, 0.05, 0.1, 0.2, 0.3)$ . It can be observed that the relative size of two confidence intervals corresponding to the same region, that is, (I), (II), (IV), (V), and (VII), is almost the same in all cases and hence rather independent of  $\sigma$ . Of course, the absolute size depends on  $\sigma$ , as can be seen from the rightmost bottom plot, where we indicate the intervals for region (II) versus different  $\sigma^2$ . The roughly linear dependence of the mean of the  $|D_T - D_c|^2$  on  $\sigma^2$  is in agreement with theory; see (20) with (22).

#### 4. FRAP Recovery Curves: Integrated Data Approach

In this section we investigate the different way of treating the data; in the following we call it the *integrated data approach*, which is a common procedure in the evaluation of FRAP experiments [1, 12]. Almost without exception, the experimental biologists are using for their purposes the so-called FRAP recovery curve, the space averaged signal  $(v(t_i))_{i=1, N_t}$ ; see below), instead of the spatiotemporal data  $u(x_i, t_i)$ . Our aim is to compare both approaches in view of sensitivity and confidence intervals.

By integrating the data, we have to take into account that the data error is integrated as well. This will in general reduce the variance and this fact causes the false opinion about the higher precision of the integrated data approach.

**Proposition 16.** *Let us consider that the full data  $u(x_i, t_i)$ ,  $l = 1, \dots, N_x$ ,  $i = 1, \dots, N_t$  are given on a rectangular spatiotemporal grid where  $N_x$  is the number of points in the ROI (either 2D or 1D) and  $N_t$  is the number of time points. The (space-)integrated data*

$$v(t_i) = \frac{1}{N_x} \sum_{l=1}^{N_x} u(x_l, t_i), \quad i = 1, \dots, N_t \quad (60)$$

*are given on a temporal grid. Assume that the data error is normally distributed and that the number of data points is large enough such that  $\text{res}^2(D_c)/(N_{\text{data}} - 1)$  is almost equal to the data error variance in either case and that the Fisher-quantile satisfies  $f_{1, N_{\text{data}}-1}(\alpha)/f_{1, N_t-1}(\alpha) \sim 1$ .*

*Then the full data approach has equal or smaller confidence interval bounds for the parameter estimates of  $D_T$ .*

*Proof.* Let us introduce for the integrated data case the parameter-to-data mapping

$$F_I(D) = v(t_i)_{i=1, N_t}. \quad (61)$$

As the spatial errors  $e_l$  are distributed according to  $e_l = \mathcal{N}(0, \sigma^2)$ , then for the integrated errors it holds that

$$e_i = \frac{1}{N_x} \sum_{l=1}^{N_x} e_l = \mathcal{N}\left(0, \frac{\sigma^2}{N_x}\right) \quad (62)$$

and we have the regression model

$$F_I(D) = v(t_i) + e_i, \quad e_i = \mathcal{N}\left(0, \frac{\sigma^2}{N_x}\right). \quad (63)$$

Note that the error variance is now reduced by a factor  $1/N_x$ .

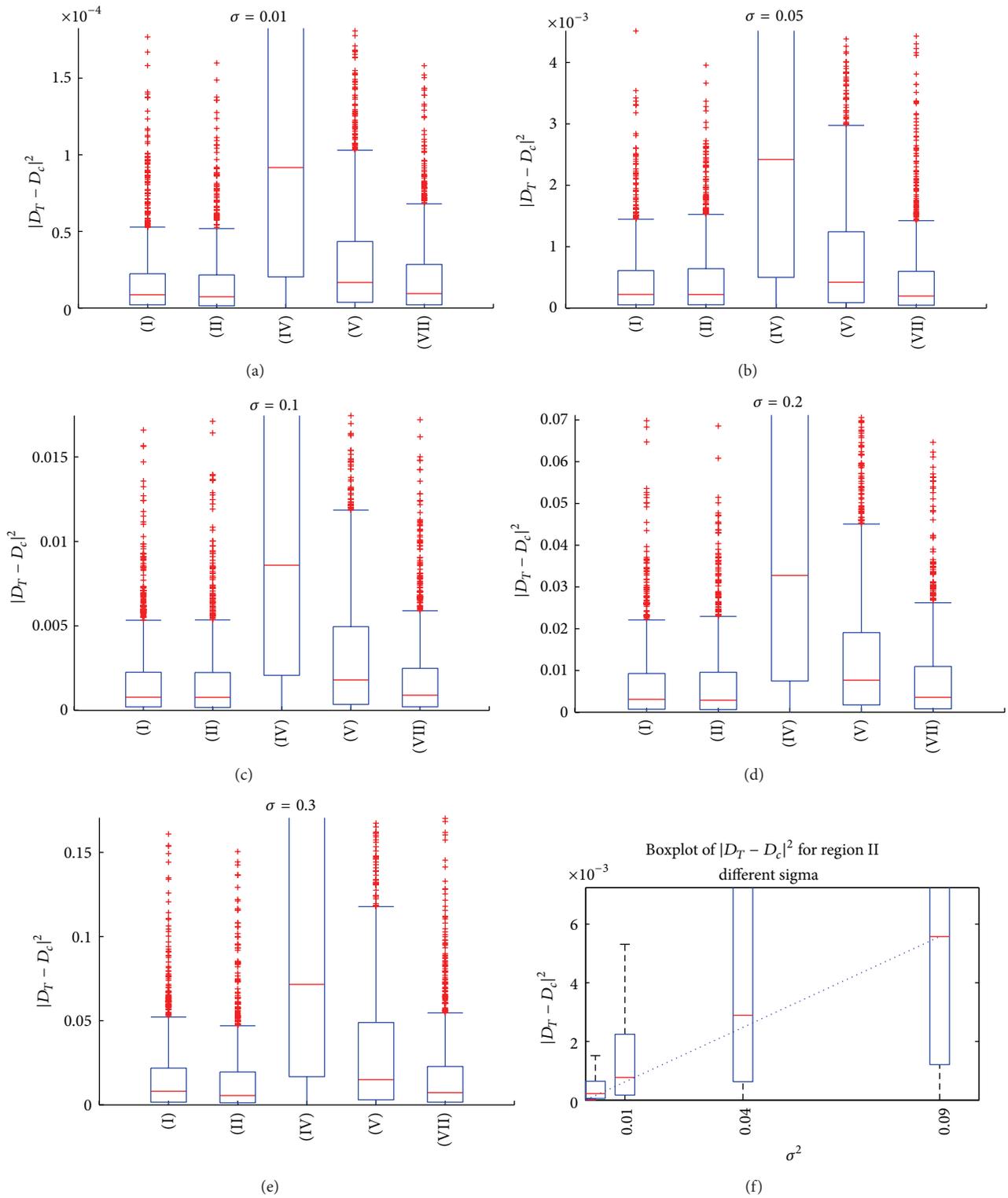


FIGURE 4: Dependence of  $|D_T - D_c|^2$  on the data noise level. Boxplot of error  $|D_T - D_c|^2$  for regions (I), (II), (IV) (V), and (VII) for 5 different  $\sigma$ . (f) shows the confidence intervals for region (II) plotted against the noise variance  $\sigma^2$ . Note: be aware of different scaling along the vertical axes of each of 6 plots.

After a regression using this model, we may employ a similar estimate for the confidence interval as before (20),

$$\begin{aligned} (D_c - D_T)^2 \sum_{i=1}^{N_t} \left( \frac{\partial}{\partial D} F_I(D) \right)^2 \\ \leq \frac{\text{res}^2(D_c)}{N_t - 1} f_{1, N_t - 1}(\alpha). \end{aligned} \tag{64}$$

Note that the sensitivity for the integrated data approach is (after changing the order of sum and derivative)

$$\sum_{i=1}^{N_t} \left( \frac{\partial}{\partial D} F_I(D) \right)^2 = \sum_{i=1}^{N_t} \left( \frac{1}{N_x} \sum_{l=1}^{N_x} \frac{\partial}{\partial D} u(x_l, t_i) \right)^2. \tag{65}$$

Compare this to the corresponding expression in the full (nonintegrated) data case, which reads

$$\sum_{i=1}^{N_t} \sum_{l=1}^{N_x} \left( \frac{\partial}{\partial D} u(x_l, t_i) \right)^2. \tag{66}$$

With the assumption that the residual error corresponds to the error variance, we find

$$\begin{aligned} \text{integrated data case: } \frac{\text{res}^2(D_c)}{N_t - 1} &\sim \frac{\sigma^2}{N_x}, \\ \text{full data case: } \frac{\text{res}^2(D_c)}{N_{\text{data}} - 1} &\sim \sigma^2. \end{aligned} \tag{67}$$

We may thus compare the sensitivities for the two approaches:

$$\begin{aligned} (D_c - D_T)_{\text{full data case}}^2 \\ \leq \frac{\sigma^2}{\sum_{i=1}^{N_t} \sum_{l=1}^{N_x} \left( \left( \frac{\partial}{\partial D} \right) u(x_l, t_i) \right)^2} f_{1, N_{\text{data}} - 1}(\alpha), \\ (D_c - D_T)_{\text{integrated data case}}^2 \leq \frac{1}{N_x} \\ \cdot \frac{\sigma^2}{\sum_{i=1}^{N_t} \left( (1/N_x) \sum_{l=1}^{N_x} (\partial/\partial D) u(x_l, t_i) \right)^2} f_{1, N_t - 1}(\alpha) \\ = \frac{\sigma^2}{\sum_{i=1}^{N_t} (1/N_x) \left( \sum_{l=1}^{N_x} (\partial/\partial D) u(x_l, t_i) \right)^2} \\ \cdot f_{1, N_t - 1}(\alpha). \end{aligned} \tag{68}$$

Now using the Cauchy-Schwarz inequality

$$\frac{1}{N_x} \left( \sum_{l=1}^{N_x} \frac{\partial}{\partial D} u(x_l, t_i) \right)^2 \leq \sum_{l=1}^{N_x} \left( \frac{\partial}{\partial D} u(x_l, t_i) \right)^2 \tag{69}$$

in the last line and the assumption  $f_{1, N_{\text{data}} - 1}(\alpha)/f_{1, N_t - 1}(\alpha) \sim 1$ , we find that the upper bound for the integrated data case is always larger than that of the full data case. Consider

$$\begin{aligned} \frac{\sigma^2}{\sum_{i=1}^{N_t} (1/N_x) \left( \sum_{l=1}^{N_x} (\partial/\partial D) u(x_l, t_i) \right)^2} f_{1, N_t - 1}(\alpha) \\ \geq \frac{\sigma^2}{\sum_{i=1}^{N_t} \sum_{l=1}^{N_x} \left( (\partial/\partial D) u(x_l, t_i) \right)^2} f_{1, N_{\text{data}} - 1}(\alpha), \end{aligned} \tag{70}$$

which finishes the proof.  $\square$

An explanation for this proposition is that although the data error variance is smaller for the integrated data case, the sensitivity is smaller as well, and this fact causes the greater or equal confidence interval bounds.

*Remark 17.* The full data case always leads to a smaller confidence interval and is thus preferable. For the one-dimensional ROI, with the uniform spatiotemporal grid

$$\begin{aligned} x_l \in \left\{ -\frac{L}{2} + \frac{(l-1)}{N_x - 1} L \mid l = 1, N_x \right\}, \\ t_i \in \{t_0 + i\Delta t \mid i = 1, N_t\}, \end{aligned} \tag{71}$$

where  $N_x$  is the number of points on the space axis,  $N_t$  is the number of time points, and  $\Delta t$  is the time step and for the reasonable case that  $0 \leq L/2r_0 \sim TD_c/r_0^2 \leq 10$ , we found the empirical estimate

$$\frac{(D_c - D_T)_{\text{full data}}^2}{(D_c - D_T)_{\text{integrated data}}^2} \leq \min \left\{ 2 \frac{r_0}{L}, 1 \right\}. \tag{72}$$

This inequality can be partially verified by analytical considerations. Indeed, the ratio on the left of (72) is roughly the ratio between the corresponding sensitivities in (69). Using a large data-size approximation, that is, approximating the sums by integrals, we can evaluate the corresponding integrals as in (43) and find

$$\begin{aligned} \frac{(D_c - D_T)_{\text{full data}}^2}{(D_c - D_T)_{\text{integrated data}}^2} \\ \sim \int_0^T \frac{(1/L) \left( \int_{-L/2}^{L/2} (\partial/\partial D) u(x, t) dx \right)^2 dt}{\int_0^T \int_{-L/2}^{L/2} |(\partial/\partial D) u(x, t)|^2 dx} \\ = \frac{\int_0^T \int_{-L/2}^{L/2} |(\partial/\partial D) u(x, t)|^2 \omega(\zeta(t)) dt}{\int_0^T \int_{-L/2}^{L/2} |(\partial/\partial D) u(x, t)|^2 dt}, \end{aligned} \tag{73}$$

with

$$\omega(\zeta(t)) = \frac{8\zeta(t)}{2\zeta(t) - 4\zeta(t)^3 + 3e^{\zeta(t)^2} \sqrt{\pi} \text{erf}(\zeta(t))}, \tag{74}$$

where  $\zeta(t) = L/\sqrt{r_0^2 + 8D_c t}$ . A plot of the function  $\omega$  reveals that it is close to 1 if  $\zeta \in [0, 1/2]$ , close to 0 if  $\zeta \geq 3$ , and

monotonically decaying from 1 to 3 for  $\zeta \in [1/2, 3]$ . Thus if  $t \in [0, T]$ , then  $\zeta(t) \in [\zeta(T), L/r_0]$ . In particular if  $L/r_0 \leq 1/2$ , then  $\omega$  in the integrand is close to 1 and the intervals are almost the same. This corresponds to the right-hand side of (72) being  $\min\{2(r_0/L), 1\} = \min\{4, 1\} = 1$ , hence verifying (72). Conversely if  $\zeta(T) \geq 3$ , then  $\omega$  is almost 0, and the ratio in (72) is close to 0. The case  $\zeta(T) \geq 3$  corresponds to  $L/r_0 \geq 3\sqrt{1 + 8D_c T/r_0^2}$ . Within our assumptions of  $L/2r_0 \sim TD_c/r_0^2$  this leads to the condition  $L/2r_0 \geq 18$  and hence the right-hand side of (72) being small ( $\sim 1/18$ ) as well.

*Remark 18.* As the proof of Proposition 16 is based on the Cauchy-Schwarz inequality, we can consider the cases in which it is sharp, that is, when the left-hand side of (69) is close to the right-hand side. This happens when the two vectors in the inequality are collinear; in our case this means that  $(\partial/\partial D)u(x_i, t_i) \sim \text{const}_i$  for all  $x_i$ . Thus, by this reasoning, if the sensitivity does not vary much in the spacial direction, then the confidence intervals in the full data and integrated data case should be approximately equal. We can roughly interpret this case as that where all data points are highly relevant and no data reduction is necessary. Conversely, if the sensitivities are not constant in spacial direction, then the full data case yields significantly smaller confidence intervals.

*Example 19.* Again, we underpin the theoretical study in this section by numerical experiments. For the plots in Figure 5, we compared the squared errors  $|D_c - D_T|^2$  for the full data case and the integrated data case for four kinds of data regions (labeled as (I)–(IV)). As data we used again a rectangular region  $(x_i, t_i) \in [-L/2, L/2] \times [0, 6]$  with different values of  $L$ . The grid spacing  $\Delta x, \Delta t$ , the value of the exact diffusivity  $D_T$ , and the noise variance  $\sigma^2$  were the same as in Examples 9 and 15. As before we used 1000 samples for each of cases (I)–(IV) and for both the full and the integrated data we computed an error distribution of  $|D_c - D_T|^2$ .

The length of the spatial interval was as follows: for case (I):  $L = 0.2$ , for case (II):  $L = 1$ , for case (III):  $L = 2$ , and for case (IV):  $L = 4$ . The time interval  $[0, 6]$  was the same for all four cases.

In the plot in Figure 5, we display four pairs of boxplots of the squared error  $|D_c - D_T|^2$  corresponding to the four cases of data regions. In each of the pairs, the left boxplot corresponds to using full data as before, while the right boxplot corresponds to using spatially integrated data (60) on the same region. The error plots are in full accordance with Proposition 16, because in each pair, the left box is smaller or equal to the right box, which means that the error using full data (left) is always smaller or equal to that of using integrated data (right). The results also fit quantitatively to the estimate in (72). Note that the factor  $L/2r_0$  is equal to 1, 5, 10, 20, for cases (I), (II), (III), and (IV). Thus, in case (I), we have  $2r_0/L = 1$  such that (72) predicts a comparable error for the full and integrated data approach. Indeed, the first pair in the boxplot has comparable sizes of the boxes. In the other three case we have  $2r_0/L < 1$ ; hence (72) predicts a larger error for the integrated data case. This can be observed again for cases (II)–(IV) where the integrated data case has a larger

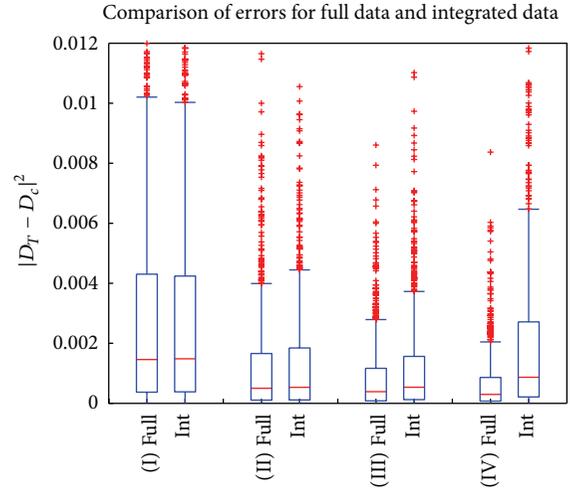


FIGURE 5: Comparison of the squared errors  $|D_c - D_T|^2$  for the full data approach and the integrated data approach for four data regions with different length of the spatial interval (the factor  $L/2r_0$  is equal to 1, 5, 10, 20, resp.) labeled as (I)–(IV); time interval was the same for all cases  $[0, 6]$ . In the pairs, the left and right boxplot correspond to using full data (Full) and spatially integrated data (Int), respectively.

error with increasing value of  $L/2r_0$ . The explanation of such a behavior resides in the inclusion of more and more larger region of irrelevant data into the spatial integration.

*Remark 20.* Unexpectedly, Figure 5 shows another interesting phenomenon, from left to right, with the growing ROI (more precisely for increasing ratio  $L/2r_0$ ); while the squared error for the full data approach monotonically decreases, there is a minimum for the integrated data approach; that is, there is an optimum (maximum sensitivity) for some value  $L/2r_0$ . Indeed, when the ROI is too small we disregard a huge amount of the relevant data, and conversely, when the ROI is too big, the integrated approach kills the valuable spatial resolution causing the loss of sensitivity of the output data to the parameter value. Hence, this is an explanation of the existence of an optimal size of ROI (i.e., an optimal value of  $L/2r_0$ ) for the integrated data approach.

## 5. Conclusion and Future Directions

In this paper, we have introduced and solved the problem of data space selection-reduction and data processing for parameter identification in a reaction-diffusion model based on FRAP experiments. First, either for a known data noise level (determined by parameter  $\tau$ ) or for previously chosen coefficient of enlargement of the confidence interval for the parameter estimate (parameter  $\eta$ ), we propose a new method to separate the relevant and irrelevant data sets.

Afterwards, we show that the data set represented by the FRAP recovery curves (the integrated data approach) leads to a larger confidence interval compared to the spatiotemporal data. This rigorous statement is worth to be promoted in

the FRAP community because it goes against the common knowledge and practice.

Four numerical examples support our conclusions. The corresponding graphical visualizations, despite their simplicity, help us in discovering new features of FRAP and FLIP techniques and inspire new investigation related to the optimal design of photobleaching experiments. For instance, the lotus flower-like plot (Figure 2) led us to the investigation related to the combination of FRAP and FLIP experiments [23]. The next step in the same direction was done in our conference paper [24], where we use the sensitivity analysis developed here to find specific conditions for optimizing the radius of bleach spot. Again, the inspiration to optimize a design factor was induced by Figure 5; that is, the existence of an optimal size of monitored region for the integrated data approach (cf. Remark 20) led us to look for an optimal value of bleach size.

Once proving the concept, our future goal is even more ambitious. Based on the same sensitivity analysis we shall suggest the optimal values for the experimental design variables, that is, the experimental protocol modifications.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the OeAD (Austrian Agency for International Mobility and Cooperation in Education, Science and Research) within the programme “Aktion Oesterreich-Tschechien (AOeCZ-Universitaetslehrerstipendien)” and by the Ministry of Education, Youth and Sports 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).

## References

- [1] F. Mueller, D. Mazza, T. J. Stasevich, and J. G. McNally, “FRAP and kinetic modeling in the analysis of nuclear protein dynamics: what do we really know?” *Current Opinion in Cell Biology*, vol. 22, no. 3, pp. 403–411, 2010.
- [2] R. Kaňa, O. Prášil, and C. W. Mullineaux, “Immobility of phycobilins in the thylakoid lumen of a cryptophyte suggests that protein diffusion in the lumen is very restricted,” *FEBS Letters*, vol. 583, no. 4, pp. 670–674, 2009.
- [3] R. Kaňa, E. Kotabová, M. Lukeš, and et al., “Phycobilisome mobility and its role in the regulation of light harvesting in red algae,” *Plant Physiology*, vol. 165, no. 4, pp. 1618–1631, 2014.
- [4] D. Axelrod, D. E. Koppel, J. Schlessinger, E. Elson, and W. W. Webb, “Mobility measurement by analysis of fluorescence photobleaching recovery kinetics,” *Biophysical Journal*, vol. 16, no. 9, pp. 1055–1069, 1976.
- [5] J. Ellenberg, E. D. Siggia, J. E. Moreira et al., “Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis,” *The Journal of Cell Biology*, vol. 138, no. 6, pp. 1193–1206, 1997.
- [6] P. Jönsson, M. P. Jonsson, J. O. Tegenfeldt, and F. Höök, “A method improving the accuracy of fluorescence recovery after photobleaching analysis,” *Biophysical Journal*, vol. 95, no. 11, pp. 5334–5348, 2008.
- [7] C. W. Mullineaux, M. J. Tobin, and G. R. Jones, “Mobility of photosynthetic complexes in thylakoid membranes,” *Nature*, vol. 390, no. 6658, pp. 421–424, 1997.
- [8] S. Seiffert and W. Oppermann, “Systematic evaluation of FRAP experiments performed in a confocal laser scanning microscope,” *Journal of Microscopy*, vol. 220, no. 1, pp. 20–30, 2005.
- [9] T.-T. Tsay and K. A. Jacobson, “Spatial fourier analysis of video photobleaching measurements: principles and optimization,” *Biophysical Journal*, vol. 60, no. 2, pp. 360–368, 1991.
- [10] O. N. Irrechukwu and M. E. Levenston, “Improved estimation of solute diffusivity through numerical analysis of FRAP experiments,” *Cellular and Molecular Bioengineering*, vol. 2, no. 1, pp. 104–117, 2009.
- [11] J. Mai, S. Trump, R. Ali et al., “Are assumptions about the model type necessary in reaction-diffusion modeling? A FRAP application,” *Biophysical Journal*, vol. 100, no. 5, pp. 1178–1188, 2011.
- [12] I. F. Szbalzarini, *Analysis, Modeling and Simulation of Diffusion Processes in Cell Biology*, VDM Dr. Muller, 2009.
- [13] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Ill-Posed Problems*, Kluwer, Dordrecht, The Netherlands, 1996.
- [14] R. Kaňa, C. Matonoha, Š. Papáček, and J. Soukup, “On estimation of diffusion coefficient based on spatio-temporal FRAP images: an inverse ill-posed problem,” in *Programs and Algorithms of Numerical Mathematics 16*, J. Chleboun, K. Segeth, J. Šístek, and T. Vejchodský, Eds., pp. 100–111, Academy of Sciences of the Czech Republic, 2013.
- [15] Š. Papáček, R. Kaňa, and C. Matonoha, “Estimation of diffusivity of phycobilisomes on thylakoid membrane based on spatio-temporal FRAP images,” *Mathematical and Computer Modelling*, vol. 57, no. 7-8, pp. 1907–1912, 2013.
- [16] K. S. Zadeh, H. J. Montas, and A. Shirmohammadi, “Identification of biomolecule mass transport and binding rate parameters in living cells by inverse modeling,” *Theoretical Biology and Medical Modelling*, vol. 3, article 36, 2006.
- [17] L. Lukšan, M. Tuma, C. Matonoha et al., “UFO 2014—interactive system for universal functional optimization,” Tech. Rep. V-1191, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, 2014, <http://www.cs.cas.cz/luksan/ufo.html>.
- [18] R. Kaňa, “Mobility of photosynthetic proteins,” *Photosynthesis Research*, vol. 116, no. 2-3, pp. 465–479, 2013.
- [19] A. Cintrón-Arias, H. T. Banks, A. Capaldi, and A. L. Lloyd, “A sensitivity matrix based methodology for inverse problem formulation,” *Journal of Inverse and Ill-posed Problems*, vol. 17, no. 6, pp. 545–564, 2009.
- [20] K. Sadegh Zadeh and H. J. Montas, “A class of exact solutions for biomacromolecular diffusion-reaction in live cells,” *Journal of Theoretical Biology*, vol. 264, no. 3, pp. 914–933, 2010.
- [21] K. Sadegh Zadeh, “A synergic simulation-optimization approach for analyzing biomolecular dynamics in living organisms,” *Computers in Biology and Medicine*, vol. 41, no. 1, pp. 24–36, 2011.
- [22] D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and Its Applications*, John Wiley & Sons, New York, NY, USA, 1988.
- [23] Š. Papáček, J. Jablonský, C. Matonoha, R. Kaňa, and S. Kindermann, “FRAP & FLIP: two sides of the same coin?” in *Proceedings of the 3rd International Conference on Bioinformatics and*

*Biomedical Engineering (IWBBIO '15), Granada, Spain, April 2015*, F. Ortuño and I. Rojas, Eds., vol. 9044, pp. 444–455, Springer, 2015.

- [24] Š. Papáček, J. Jablonský, R. Kaňa, C. Matonoha, and S. Kindermann, “From data processing to experimental design and back again: a parameter identification problem based on FRAP images,” in *Proceedings of the 17th International Conference on Digital Image Processing (ICDIP '15)*, Dubai, UAE, January 2015.