

O některých optimálních parametrech pro FRAP experimenty

Ctirad Matonoha

Ústav informatiky AV ČR, Praha

PANM 18, 19. - 24. 6. 2016, Janov nad Nisou

Společně s:

R. Kaňa (MBÚ Třeboň)

Š. Papáček (FROV JČU České Budějovice)

J. Jablonský (FROV JČU České Budějovice)

S. Kindermann (JKU Linz)

B. Macdonald (University of Glasgow)

O některých optimálních parametrech pro FRAP experimenty (On Model-Based Design of Photobleaching Experiments)

Zaměříme se na to, jak určit některé parametry pro FRAP experimenty k dosažení co nejlepšího řešení. V této oblasti se dosáhlo několik zajímavých a pro praxi užitečných výsledků. Např.:

- Špatná podmíněnost problému když difuze závisí na čase
- Ekvivalence problémů, kde lze využít znalosti úrovně šumu v datech
- Tvar lotosového květu pro výběr relevantních dat
- Existence optimálního poloměru záblesku
- Disk vs. anulus jako vítězové optimálního tvaru záblesku

- 1 Motivace
- 2 Matematický model
- 3 Určení difuzního koeficientu
- 4 Citlivostní analýza

- 1 Motivace
- 2 Matematický model
- 3 Určení difuzního koeficientu
- 4 Citlivostní analýza

FRAP (Fluorescence Recovery After Photobleaching = Fluorescenční zotavení po fotozáblesku) technika umožňuje určit difuzi fluorescenčních objektů (bílkovin).

Tato metoda je založena na aplikaci krátkého silného laserového ozáření (tzv. záblesk) na malou cílovou oblast buňky která způsobuje nevratnou ztrátu ve fluorescenci v této oblasti.

Po aplikaci záblesku je pozorováno zotavení ve fluorescenci v dané oblasti, které vyjadřuje difuzi fluorescenčních objektů z prostoru vně záblesku.

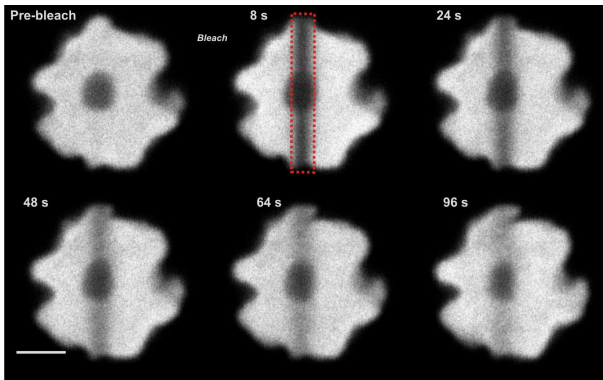


R. Kaňa, E. Kotabová, M. Lukeš, Š. Papáček, C. Matonoha, L.N. Liu, O. Prášil, C.W. Mullineaux

Phycobilisome Mobility and Its Role in the Regulation of Light Harvesting in Red Algae.

Plant Physiology, Vol. 165, 2014, p. 1618-1631.

Experimentální FRAP data

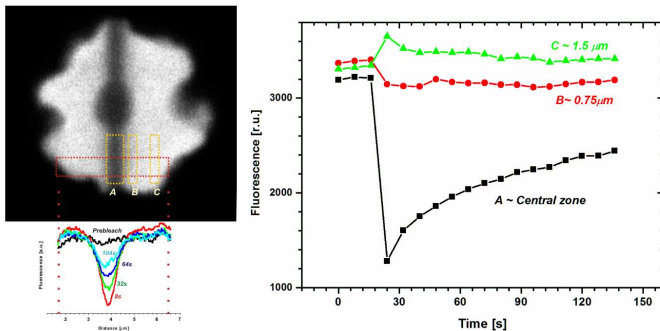


Experimentální posloupnost dat získaných technikou FRAP pro buňku řasy *Porphyridium cruentum*.

Pre-bleach: fluorescenční obrázek před laserovým zábleskem.

Poté proveden kruhový záblesk přes střed buňky ve vertikálním směru (červený obdélník) a pozorován návrat fluorescence.

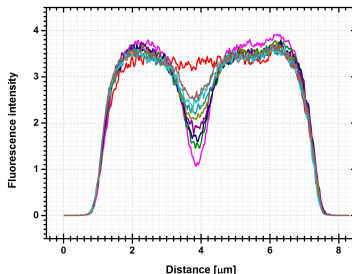
Zpracování experimentálních dat



Vlevo: Typický obrázek 8s po záblesku. Lze pozorovat různé oblasti, centrální i mimo (žluté obdélníky). Dole je dynamika zotavení v centrální oblasti v závislosti na čase. Křivky byly zkonstruovány na základě 2D obrázků zprůměrováním dat podél os kolmých na pruh záblesku.

Vpravo: Zotavovací křivky jako časová závislost prostorově zprůměrovaných hodnot fluorescence v jednotlivých oblastech.

Prostorově-časový obrázek experimentálních dat



Experimentální (zašuměná) data ve formě 1D profilů pro různé časové okamžiky po záblesku. Červená křivka je situace před zábleskem.

Souřadnice x: Pozice podél osy kolmé na pruh záblesku.

Souřadnice y: Odpovídající zprůměrovaná hodnota fluorescence.

Uprostřed je vidět postupné zotavení signálu od nejnižší hodnoty (první po záblesku) po nejvyšší (ustálený stav).

- 1 Motivace
- 2 Matematický model**
- 3 Určení difuzního koeficientu
- 4 Citlivostní analýza

Inverzní problém

Obecný tvar difuzní rovnice:

$$\frac{\partial y}{\partial t} - D \frac{\partial^2 y}{\partial x^2} = 0, \quad x \in [a, b], \quad t \in [0, T]. \quad (1)$$

Počáteční podmínka

$$y(x, 0) = f(x), \quad x \in [a, b]. \quad (2)$$

Okrajová podmínka buď Dirichletova typu

$$y(a, t) = g_a(t), \quad y(b, t) = g_b(t), \quad t \geq 0, \quad (3)$$

nebo Neumannova typu

$$-D \frac{\partial y}{\partial x}(a, t) = h_a(t), \quad D \frac{\partial y}{\partial x}(b, t) = h_b(t), \quad t \geq 0. \quad (4)$$

Inverzní problém: Na základě vstupních dat $y(x, t)$ nás zajímá $D \geq 0$.

Matice experimentálních dat

Na základě FRAP experimentů máme matici dat

$$y_{exp}(x_i, t_j), \quad i = 1 \dots N, \quad j = -m \dots M,$$

kde jsou provedena měření:

- N je počet prostorových bodů
- m je počet hodnot před zábleskem
- $M + 1$ je počet hodnot po záblesku (počet časových řad)

Následně,

$$x_1 = a, \quad x_N = b,$$

$t_0 = 0$ odpovídá prvnímu měření po záblesku a tedy

- $y_{exp}(x_i, t_0)$, $i = 1 \dots N$, reprezentuje počáteční podmínku
- $y_{exp}(a, t_j)$, $j = 0 \dots M$, repr. levou Dirichletovu okrajovou podmínku
- $y_{exp}(b, t_j)$, $j = 0 \dots M$, repr. pravou Dirichletovu okrajovou podmínku
- Neumannova okrajová podmínka pro každé j je určena použitím prvního Fickova zákona

Testovací data

Pro teoretické vyšetřování, porovnání metod a další analýzu zvolíme libovolné $D \equiv D_{true}$ a vyrobíme testovací data, gaussovky

$$g(x_i, t_j) = \frac{y_{0,0} r_0}{\sqrt{r_0^2 + 8Dt_j}} \exp \frac{-2x_i^2}{r_0^2 + 8Dt_j}, \quad x_i \in [-\frac{L}{2}, \frac{L}{2}], \quad t_j \in [0, T],$$

které jsou řešením PDE (1)–(3). Zde

- $y_{0,0}$ je výška křivky, odpovídá prvnímu měření po záblesku (počáteční podmínka v čase $t_0 = 0$), lze volit $y_{0,0} = 1$
- r_0 je pološířka (poloměr záblesku), platí $g(r_0, 0) = y_{0,0} \exp(-2)$

Testovací zašuměná data:

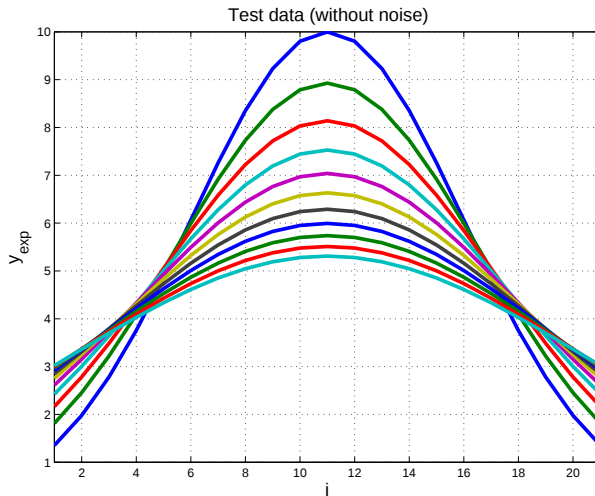
$$y_{test}(x_i, t_j) = g(x_i, t_j) + e_{ij},$$

kde e_{ij} je aditivní gaussovský šum,

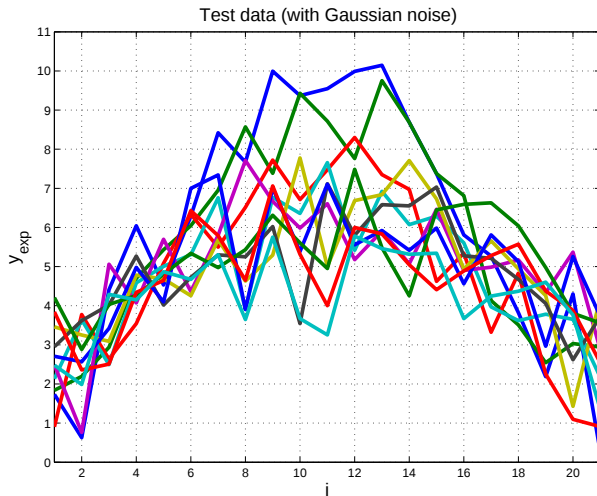
$$\sum_{i=1}^N e_{ij} = 0, \quad \sum_{i=1}^N e_{ij}^2 = N\sigma^2, \quad j = 0 \dots M,$$

a σ je úroveň šumu. BÚNO $y_{exp} \equiv y_{test} \equiv y$.

Testovací data bez šumu



Testovací data zašuměná



- 1 Motivace
- 2 Matematický model
- 3 Určení difuzního koeficientu
- 4 Citlivostní analýza

$$z^{Mull}(t_j) = y(0, t_j) = \frac{y_{00} r_0}{\sqrt{r_0^2 + 8D_j^{Mull} t_j}}$$

kde $z^{Mull}(t_j)$ jsou naměřená (nebo testovací) data v centru monitorovací oblasti ($x = 0$). Odtud spočítáme D_j^{Mull} , $j = 1 \dots M$, a jejich průměrnou hodnotu D^{Mull} .

 C.W. Moullineaux, M.J. Tobin, G.R. Jones

Mobility of photosynthetic complexes in thylakoid membranes.

Nature, 390:421–424, 1997.

Hypotéza: Nezřídka se může stát, že v závislosti na σ a přesném D_{true} jde spočtené D^{Mull} do nekonečna.

$$z^{Elle}(t_j) = 1 - \frac{I_{t_j}}{I_{t_0}}, \quad I_{t_j} = \int_{-r_0}^{r_0} y(x_i, t_j) dx$$

$$D_j^{Elle} = \frac{r_0^2}{\pi t_j} \frac{1 - [1 - z^{Elle}(t_j)]^2}{[1 - z^{Elle}(t_j)]^2} = \frac{r_0^2}{\pi t_j} \frac{I_{t_0}^2 - I_{t_j}^2}{I_{t_j}^2}$$

pro $j = 1 \dots M$ a D^{Elle} je jejich průměrná hodnota.



J. Ellenberg, E.D. Siggia, J.E. Moreira, C.L. Smith, J.F. Presley, H.J. Worman, J. Lippincott-Schwartz

Nuclear membrane dynamics and reassembly in living cells: targeting of an inner nuclear membrane protein in interphase and mitosis.

The Journal of Cell Biology, 138:1193–1206, 1997.

Hypotéza: I když spočtené D^{Elle} v podstatě nezávisí na σ , může být hodně daleko od přesného D_{true} .

Určení difuzního koeficientu - PDE (Papáček et.al.)

Zkonstruujeme **objektivní funkci** Y reprezentující rozdíl mezi experimentálními (testovacími) a simulovanými profily a najdeme takovou hodnotu D , která minimalizuje Y . Obvyklým tvarem je součet čtverců

$$Y(D) = \sum_{j=1}^M \sum_{i=1}^N [y(x_i, t_j) - y_{sim}(x_i, t_j, D)]^2, \quad (5)$$

kde $y_{sim}(x_i, t_j, D)$ jsou numerické hodnoty splňující (1)–(3), a řešíme problém

$$D^{PDE} = \arg \min_D Y(D).$$



Š. Papáček, R. Kaňa, C. Matonoha

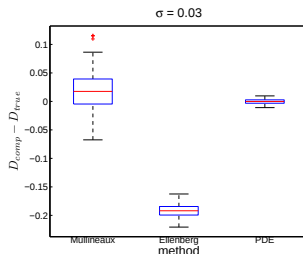
Estimation of diffusivity of phycobilisomes on thylakoid membrane based on spatio-temporal FRAP images.

Mathematical and Computer Modelling, 57: 1907–1912, 2013.

Otázka: Jak zvolit okrajovou podmínku?

Výsledky pro testovací zašuměná data

Vygenerováno 1000 datových sad s různými šумы pro hodnotu $\sigma = 0.03$, D_{comp} značí D^{Mull} , D^{Ele} , D^{PDE} .



Detailní porovnání metod, viz



Š. Papáček, B. Macdonald, C. Matonoha

Closed-form formulae vs. PDE based numerical solution for the FRAP data processing: Theoretical and practical comparison.

Submitted.

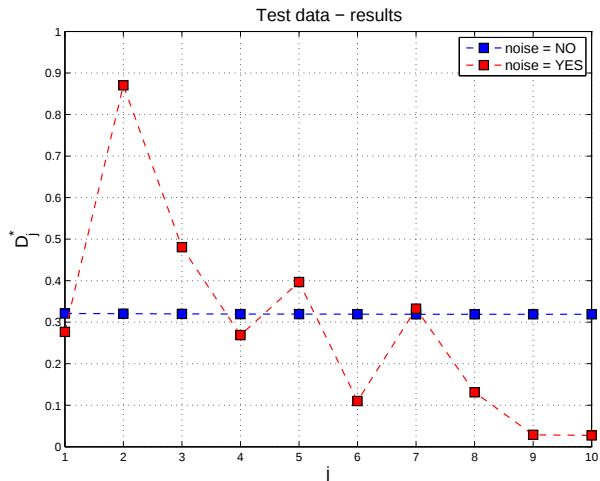
Vezmeme-li v úvahu biologickou realitu spočívající v možné závislosti difuze na čase, pak řešíme problém

$$D_j^* = \arg \min_{D_j} Y_j(D_j),$$

kde

$$Y_j(D_j) = \sum_{i=1}^N [y(x_i, t_j) - y_{sim}(x_i, t_j, D_j)]^2, \quad j = 1, \dots, M.$$

Výsledky pro časově závislou difuzi



Tikhonovova regularizace

Toto je špatně podmíněná úloha v tom smyslu, že řešení, tj. difuzní koeficienty $D_1^*, \dots, D_M^*$ nezávisí spojitě na datech. Musí se uvažovat regularizované řešení

$$D_j^*(\alpha) = \arg \min_{D_j} Y_j(D_j, D_{reg}, \alpha),$$

kde

$$Y_j(D_j, D_{reg}, \alpha) = \sum_{i=0}^N [y_{exp}(x_i, t_j) - y_{sim}(x_i, t_j, D_j)]^2 + \alpha (D_j - D_{reg})^2$$

pro $j = 1 \dots M$, kde $\alpha \geq 0$ je regularizační parametr a D_{reg} je očekávaná hodnota.

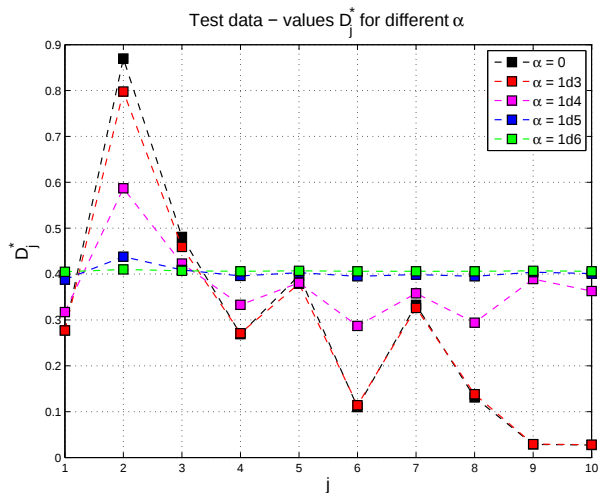


Š. Papáček, R. Kaňa, C. Matonoha

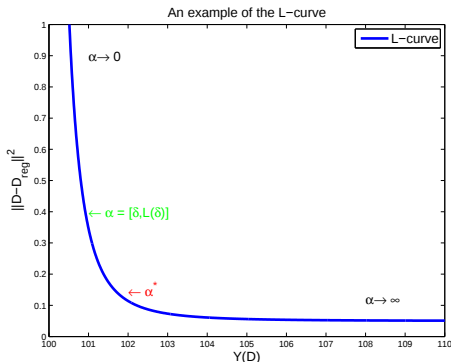
Estimation of diffusivity of phycobilisomes on thylakoid membrane based on spatio-temporal FRAP images.

Mathematical and Computer Modelling, 57: 1907–1912, 2013.

Výsledky pro časově závislou difuzi - regularizace



Jak určit optimální α^* ? Pomocí L-křivky:



Per Christian Hansen

Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion.

SIAM, 1997.

Ekvivalentní problémy

Problém

$$D_j^*(\alpha) = \arg \min_{D_j} Y_j(D_j, D_{\text{reg}}, \alpha) \quad (6)$$

je ekvivalentní následujícím optimalizačním problémům

$$D^*(\delta) = \arg \min_D \|D - D_{\text{reg}}\|^2, \quad \text{s.t.} \quad Y(D) \leq \delta, \quad D_j \geq 0 \quad (7)$$

a

$$D^*(\delta) = \arg \min_D Y(D), \quad \text{s.t.} \quad \|D - D_{\text{reg}}\|^2 \leq L(\delta), \quad D_j \geq 0 \quad (8)$$

kde

$$Y(D) = \sum_{j=1}^M Y_j(D_j)$$

s

$$D = (D_1, \dots, D_M)^T \in \mathcal{R}^M$$

a δ je funkcí σ .

Lze dokázat, že všechny tři problémy (6)–(8) jsou ekvivalentní, neboť KKT podmínky optimality vedou na stejný systém rovnic. L-křivka je spojitá a klesající, minima je dosaženo na hranici oblasti a platí

$$Y(D) = \delta \quad \Leftrightarrow \quad \|D - D_{\text{reg}}\|^2 = L(\delta).$$

Protože nalézt optimální α^* je obtížné, lze při znalosti úrovně šumu σ (nebo jeho odhadu) řešit problém (7) nebo (8) namísto problému (6).



C. Matonoha, Š. Papáček

On the connection and equivalence of two methods for solving an ill-posed inverse problem based on FRAP data.

J. Comput. Appl. Math., 290 (2015), 598–608.

Všechny minimalizační úlohy byly řešeny pomocí systému UFO



L. Lukšan, M. Tůma, C. Matonoha, J. Vlček, N. Ramešová, M. Šiška,
J. Hartman

UFO 2014 - Interactive system for universal functional optimization.

TR V-1218, ICS AS CR, Prague, 2014, 388 p.

<http://www.cs.cas.cz/luksan/ufo.html>

- 1 Motivace
- 2 Matematický model
- 3 Určení difuzního koeficientu
- 4 Citlivostní analýza**

Konfidenční interval

Na základě výsledků v knize



D.M. Bates, D.G. Watts

Nonlinear regression analysis and its applications.

John Wiley & Sons, Inc., New York, 1988.

je možné určit chybu mezi spočtenou hodnotou D_{comp} a skutečnou hodnotou D_{true} :

$$(D_c - D_t)^2 \sum_{k=1}^{N_{data}} \left[\frac{\partial}{\partial D} y(x_k, t_k) \right]^2 \leq \approx \sigma^2 \quad (9)$$

kde $N_{data} = N \times M$ je celkový počet bodů, $(D_c - D_t)^2$ je konfidenční interval a $\sum_{i=k}^{N_{data}} \left[\frac{\partial}{\partial D} y(x_k, t_k) \right]^2$ je Fisherova informační matice (FIM).

Platí, že čím větší je FIM, tím menší je horní odhad na konfidenční interval (rozdíl přesného a spočteného řešení). Odhad (9) lze prakticky využít několika způsoby.

Předpokládejme, že budeme pracovat s méně daty, tzv. redukovanými daty

$$y(\bar{x}_k, \bar{t}_k)_{k=1}^{\bar{N}_{\text{data}}} \in \mathcal{R}^{\bar{N}_{\text{data}}},$$

kde

$$\bar{N}_{\text{data}} < N_{\text{data}} \quad \text{a} \quad \{(\bar{x}_k, \bar{t}_k)\} \subset \{(x_k, t_k)\},$$

a spočítáme $\bar{D}_c$ použitím těchto dat.

Nechť $\{(\bar{x}_k, \bar{t}_k)_{k=1}^{\bar{N}_{\text{data}}}\} = \{(x_k, t_k)_{k=1}^{N_{\text{data}}}\}$. Pak pro $\bar{D}_c$ platí podobný, ale větší odhad

$$(\bar{D}_c - D_T)^2 \sum_{k=1}^{N_{\text{data}}} \left[\frac{\partial}{\partial D} y(x_k, t_k) \right]^2 \leq \approx (1 + \eta) \sigma^2. \quad (10)$$

Význam koeficientu η stanovuje následující lemma.

Lemma

Nechť $\bar{N}_{\text{data}}$ je počet redukovaných (relevantních) bodů a necht' η je zvoleno tak, že

$$\eta \geq \frac{\sum_{k=\bar{N}_{\text{data}}+1}^{N_{\text{data}}} \left(\frac{\partial}{\partial D} y(x_k, t_k) \right)^2}{\sum_{k=1}^{\bar{N}_{\text{data}}} \left(\frac{\partial}{\partial D} y(x_k, t_k) \right)^2}. \quad (11)$$

Pak pro $\bar{D}_c$ platí konfidenční interval (10).

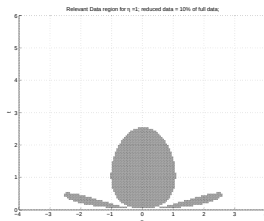


S. Kindermann, Š. Papáček

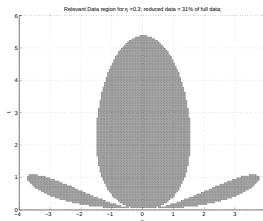
On data space selection and data processing for parameter identification in a reaction-diffusion model based on FRAP experiments.

Abstract and Applied Analysis, Art. ID 859849, 2015, 17 p.

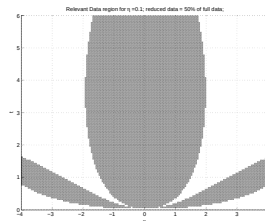
η -relevantní data



$$\eta = 1, \frac{\bar{N}_{\text{data}}}{N_{\text{data}}} = 10\%$$



$$\eta = 0.3, \frac{\bar{N}_{\text{data}}}{N_{\text{data}}} = 31\%$$



$$\eta = 0.1, \frac{\bar{N}_{\text{data}}}{N_{\text{data}}} = 50\%$$

Šedé oblasti jsou relevantní data ($\bar{N}_{\text{data}}$), do hodnoty FIM přispívají víc (větší citlivost) než irelevantní data (bílé oblasti). Poměr $\frac{\bar{N}_{\text{data}}}{N_{\text{data}}}$ indikuje η -relevantní data. Číslo $1 + \eta$ značí zvětšení konfidenčního intervalu pro $\bar{D}_c$ ve srovnání s D_c .



Š. Papáček, J. Jablonský, C. Matonoha, R. Kaňa, S. Kindermann
FRAP & FLIP: Two sides of the same coin?

In: F. Ortuno and I. Rojas (Eds.), IWBBIO 2015, Part II, p. 444-455.

Optimální poloměr záblesku

Maximalizací citlivosti (tj. maximalizací hodnoty FIM a tím minimalizací odhadu konfidenčního intervalu) lze optimalizovat některé experimentální parametry, např. poloměr záblesku r_0 .

Uvažujme případ husté sítě pozorování na prostorově-časovém cylindru $Q = [-\frac{L}{2}, \frac{L}{2}] \times [0, T]$ v 1D případě a zkusme najít optimální poloměr záblesku r_0^{opt} , který vede na největší citlivost.

FIM definujeme jako funkci r_0

$$S(r_0) = \sum_{k=1}^{N_{\text{data}}} \left[\frac{\partial}{\partial D} y(x_k, t_k) \right]^2$$

a najdeme její maximální hodnotu

$$S(r_0^{opt}) = \max_{r_0 > 0} S(r_0).$$

Analyticky: sumu aproximujeme pomocí integrálu a dojdeme k výsledku

$$r_0^{opt} \approx 1.728\sqrt{TD_c}. \quad (12)$$

Numericky: viz



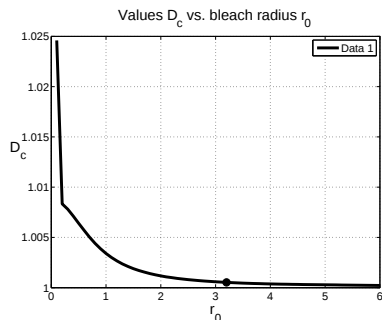
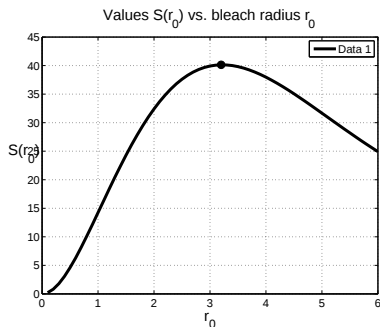
C. Matonoha, Š. Papáček

Parameter identification problem based on FRAP images: From data processing to optimal design of photobleaching experiments.

T. Kozubek et al. (Eds.): HPCSE 2015, LNCS 9611, pp. 1–10, 2016.

Optimální poloměr záblesku - výsledky

Data set	D_t	T	r_0^{opt}	D_c	$1.728\sqrt{TD_c}$
Data 1	1	4	3.2	1.000528	3.457
Data 2	2	2	3.2	2.001091	3.457
Data 3	1	2	2.4	1.000810	2.445



Optimalizace počáteční podmínky

Zkusme nyní najít takovou počáteční podmínku $y_0(x) = y(x, 0)$, aby hodnota FIM byla maximální a tedy očekávaná chyba ve spočteném parametru D byla minimální.

Zafixujme hodnoty N (počet prostorových bodů), M (počet časových řad), $x \in [-R, R]$, $t \in [0, T]$. Jedinou možností jak maximalizovat hodnotu FIM je uvažovat počáteční podmínku $y_0(x)$ jako parametr.

Ve FRAP experimentech je obvyklé, že $y_0(x)$ je binární funkce s pevnou kladnou a nulovou hodnotou. BÚNO můžeme předpokládat, že $y_0(x)$ je $\{1, 0\}$ -funkce. Řešíme tedy problém

$$\max_{B \subset \mathcal{R}^2} \sum_{k=1}^{N_{\text{data}}} \left[\frac{\partial}{\partial D} y(x_k, t_k) \right]^2 \quad \text{s.t.} \quad y_0(x) = \begin{cases} 1 & x \in B \\ 0 & \text{jinde} \end{cases}$$

Disk vs. anulus

Pro jednoduchost budeme uvažovat tvary disk, anulus, disk+anulus, dvojitý anulus, tj.

$$\text{disk: } y_0(x) = (\mathbf{1}, \mathbf{0})$$

$$\text{anulus: } y_0(x) = (\mathbf{0}, \mathbf{1}, \mathbf{0})$$

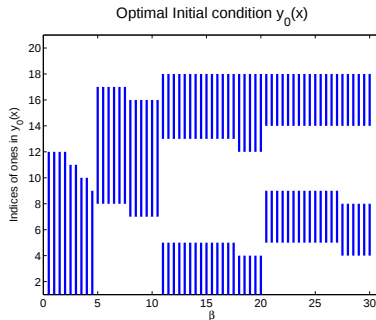
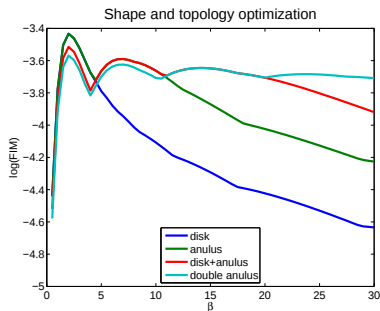
$$\text{disk+anulus: } y_0(x) = (\mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{0})$$

$$\text{dvojitý anulus: } y_0(x) = (\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{1}, \mathbf{0})$$

Definujeme

$$\beta = \frac{R^2}{4TD} \equiv \text{const} \cdot \frac{1}{D}$$

Výsledky optimalizace



Vlevo: Optimální konfigurace pro různé difuzní koeficienty.

Vpravo: Optimální tvary počáteční podmínky $y_0(x)$ (pro $N = 21$).

Hypotéza: Pro zvětšující se hodnoty β jsou optimální složitější tvary.

Otázka: Je technicky možné, aby měl záblesk tvar anulusu?

Otázka: Jak získat hodnotu β předem aby se mohl použít optimální tvar?

Citlivostní analýza a následná maximalizace FIM je základem pro optimalizaci FRAP experimentů.

Výsledky založené na citlivostní analýze se dají použít k určení relevantních dat a k volbě optimálních parametrů jako jsou poloměr a tvar záblesku.

Naším cílem je prosadit získané výsledky mezi FRAP komunitu (není moc vítáno) a doporučujeme je zohlednit při praktických měřeních.