

Bayesian information-theoretic approach to determine effective scanning protocols of cancer patients

Heyrim Cho

Assistant Professor

School of Mathematical and Statistical Sciences
Arizona State University (ASU)

May 2025

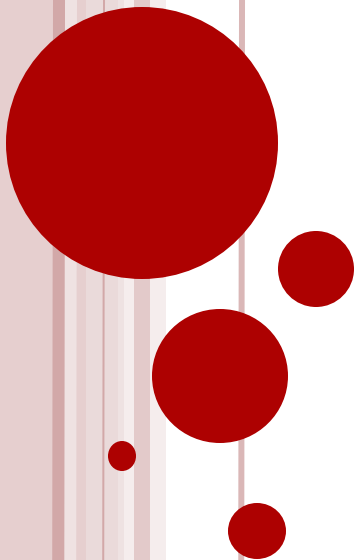
ICERM

Joint work with

Dr. Allison Lewis (Lafayette Univ.)

Dr. Kathleen Storey (Univ. Minnesota)

Dr. Tin Phan (LANL)



Cho lab - Mathematics in Biomedical Application

1. Develop mathematical models of novel medical treatments

- Cell-based immunotherapy (adoptive T-cell)
- Neural stem cell treatments

2. Develop mathematical models using new biomedical data acquisition technologies.

- Single-cell gene sequencing data
- Spatiotemporal transcriptomic data

3. Develop framework to bring mathematical models to clinical practices

- Personalized patient scanning schedule
- Design experiments to infer cancer cell interaction

Clinical collaborators :



Content

- Motivation and Problem set up
- Bayesian information-theoretic framework to determine patient scanning protocols
 - Mutual information
 - Mutual information with temporal penalty
 - Nonlinear mixture model for prior estimates
- Summary and Future work

Joint work with

Allison Lewis (Lafayette College)

Katie Storey (University of Minnesota)

Tin Phan (LANL)



Motivating questions

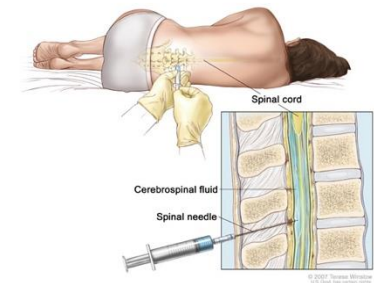
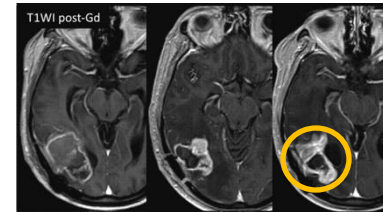
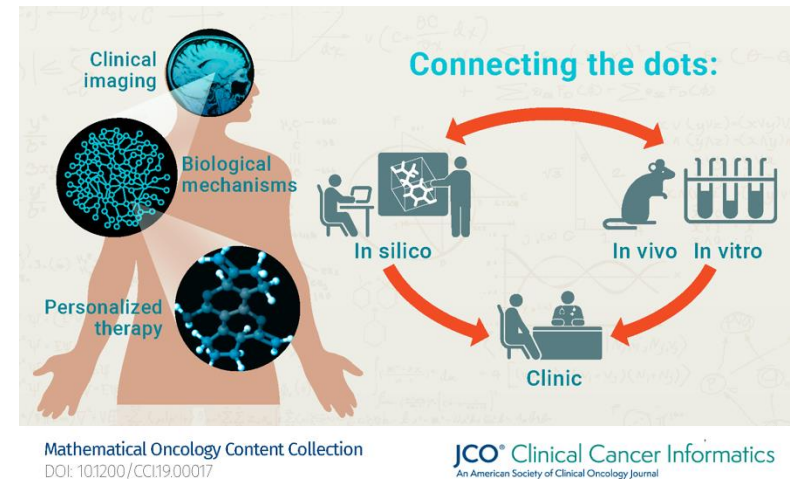
GOAL: Develop framework in which mathematicians can support decision-making in the clinic.

e.g. use mathematical models to predict the tumor growth and response to treatment of patients

- Observable cancer patient data

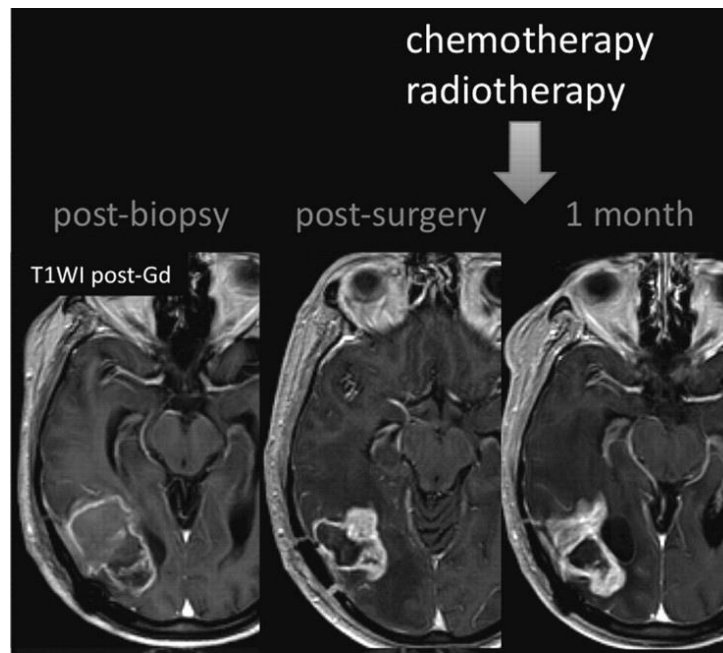
- Imaging (MRI/CT/PET)
 - : Tumor volume, Necrotic fraction, tissue structure, metabolism, ...
- Tumor Biopsy, Cerebrospinal fluid (CSF) collection, Bone marrow biopsy, peripheral blood draw
 - : Tumor composition, Genetics, Immune cell counts, Tumor antigens
- Single-cell gene sequencing data
 - : genetic profile, genetic mutation

→ **Problem: collecting clinical data can be expensive and/or invasive**



Limited temporal aspect is a challenge in model calibration using clinical data

For example, during 6-8 weeks of radiotherapy, patients receive 1-2 scan before treatment, and one more scan after 6-8 weeks treatment.



L.C. Hygino da Cruz, et al. 2011

***MRI-LINAC**



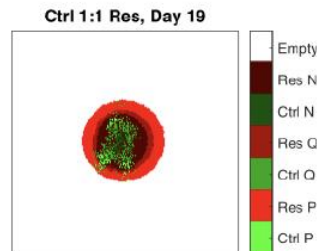
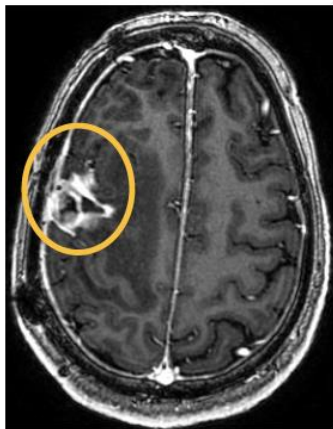
Questions:

Can we predict the treatment response earlier?

Can we determine the scan schedule in some optimal matter?

A simple set up of Model Calibration to Clinical Data

- Data $\{d_i\}$
- CT or MRI detectable tumor volume



* Note. We used synthetic data - hybrid Cellular Automata (CA) that track tumor growth in space with cell cycle and oxygen

- Model $\{Y(t_i; \theta)\}$
- Consider a dynamical system that tracks tumor volume $Y(t)$ in time

$$\frac{dY}{dt} = \lambda Y \left(1 - \frac{Y}{K}\right) - \underbrace{(1 - e^{-\alpha d - \beta d^2})u(t)Y}_{\text{cell kill due to radiotherapy}},$$

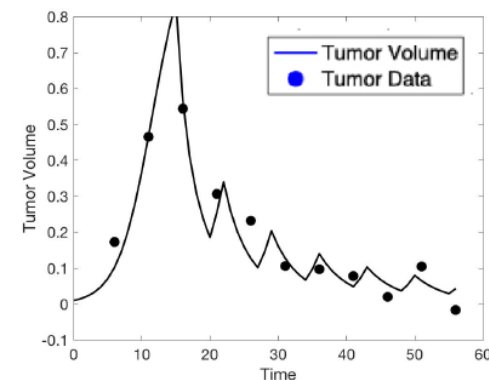
$$\theta = [\lambda, K, \alpha, \beta]$$

α, β : radiosensitivity
 λ, K : growth and capacity

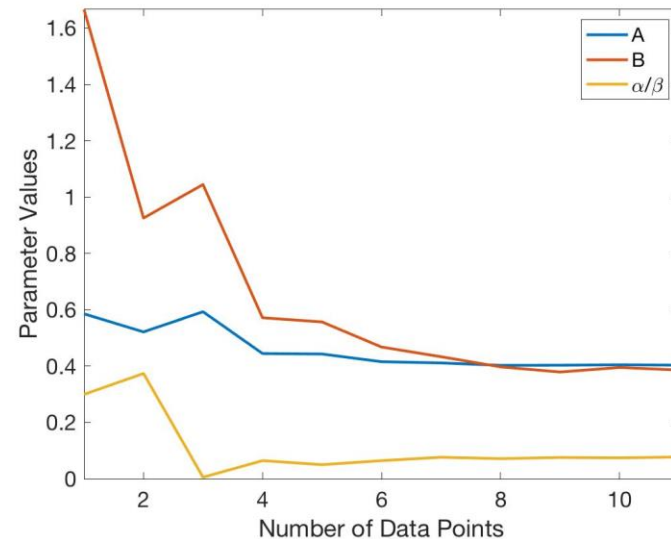
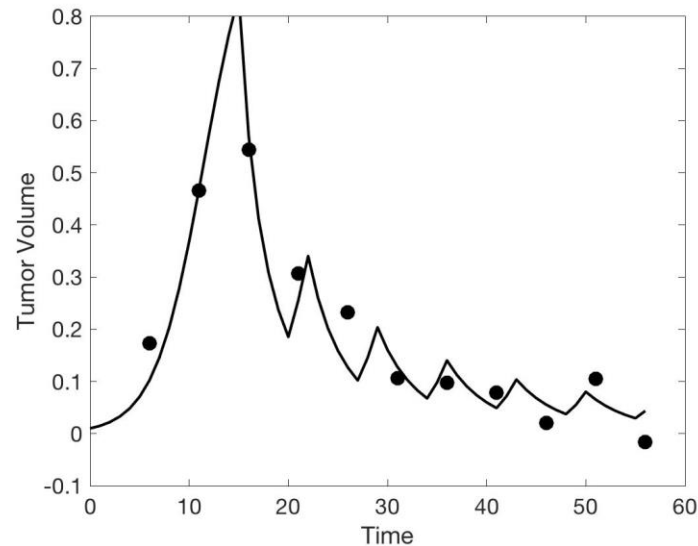
→ Model calibration using data

$$\theta^* = \operatorname{argmin}_{\theta} \|\vec{d} - Y(\vec{t}; \theta)\|_2^2$$

$$p_{\text{post}}(\theta|d) = \frac{p(d|\theta) p_{\text{prior}}(\theta)}{p(d)}$$



Sequential model calibration procedure



: Parameter values “settle down” around the 4th data point

GOAL. Determine a patient specific data collection strategy (scanning schedule) that calibrates the model parameters **early** and **accurately** using **limited number of scan budget**?

Bayesian experimental design

- Experimental design problem

$$x^* = \operatorname{argmax}_{x \in \Xi} f(x)$$

where

- x represents the design or scenario of an experiment
- Ξ is all the possible designs or scenarios
- f is an objective function that quantifies the goal of the experiment

- Approaches

$$x_n = \operatorname{argmax}_{x \in \Xi} u(x; x_{1:n-1}, f_{1:n-1})$$

where u is the acquisition function.

- Model-based approach
 - Information-theoretic approach
 - Mean objective cost of uncertainty
- Data-driven approach (Gaussian process)
 - expected improvement (EI), probability of improvement (PI), upper-confidence bounds (UCB), entropy search (ES), ...

PART 1. Mutual Information

Application to Prostate Cancer with Radiotherapy

Bayesian Information-theoretic Calibration

⑩ Bayesian calibration framework d : data θ : parameter

$$p_{\text{post}}(\theta|d) = \frac{p(d|\theta) p_{\text{prior}}(\theta)}{p(d)}$$



Likelihood $p(d|\theta)$: Gaussian distribution

Estimate posterior using MCMC :

DRAM (Delayed Rejection Adaptive Metropolis Algorithm)

H. Haario et al. (2006) DRAM: Efficient adaptive MCMC, Stats. Comput.

- Sequential calibration : Given a data set of size $n - 1$ (D_{n-1}), choose the n -th data d_n in an optimal manner.

$$d_n = \operatorname{argmax}_{d \in \Xi} u(d; D_{n-1}) \text{ for some acquisition function } u .$$

→ Take u as the reduction of model uncertainty via **mutual information**

Mutual Information Framework

Goal: Maximize Mutual Information between model parameter θ and future data d_n . (D_{n-1} is previously chosen $n - 1$ scans)

Shannon Entropy:

$$H(\Theta|D_{n-1}) = - \int_{\Omega} p(\theta|D_{n-1}) \log(p(\theta|D_{n-1})) d\theta$$

Utility Function:

$$U(d_n, \xi_n) = \int_{\Omega} p(\theta|d_n, D_{n-1}) \log p(\theta|d_n, D_{n-1}) d\theta - \int_{\Omega} p(\theta|D_{n-1}) \log p(\theta|D_{n-1}) d\theta$$

Mutual Information:

$$\begin{aligned} I(\theta; d_n | D_{n-1}, \xi_n) &= \int_{\mathcal{D}} U(d_n, \xi_n) p(d_n | D_{n-1}, \xi_n) dd_n \\ &= \int_{\mathcal{D}} \int_{\Omega} p(\theta, d_n | D_{n-1}, \xi_n) \log \frac{p(\theta, d_n | D_{n-1}, \xi_n)}{p(\theta | D_{n-1}) p(d_n | D_{n-1}, \xi_n)} d\theta dd_n \quad (1) \end{aligned}$$

Given D_{n-1} , choose the n -th scan such that

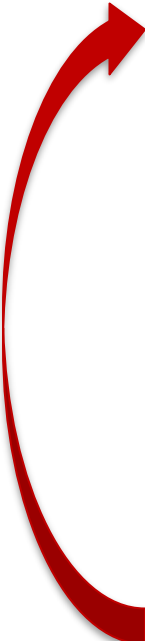
$$\xi_n^* = \arg \max_{\xi_n \in \Xi} I(\theta; d_n | D_{n-1}, \xi_n)$$

A. Kraskov et al., Estimating mutual information, PRE 69 (2004)
A Lewis, R Smith, et al. JCP (2018)

Algorithm – multi-fidelity simulation

d_ℓ Low-fidelity (in-silico, ODE) model prediction at design n

\tilde{d}_n High-fidelity (experimental, patient) data collected at design n



Existing data: $D_{n-1} = \{\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{n-1}\}$

↓ Run MCMC (DRAM) Algorithm

Calibrate parameters of low-fidelity model: $d_\ell(\theta, \xi_n)$

↓ Compute Mutual Information

Choose new design $\xi_n^* = \arg \max_{\xi_n \in \Xi} I(\theta; d_n | D_{n-1}, \xi_n)$ to reduce uncertainty in θ

↓ Collect High-Fidelity data

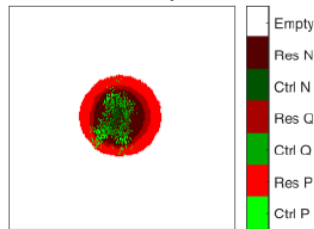
Add $\tilde{d}_n = d_h(\xi_n^*) + \tilde{\varepsilon}_n(\xi_n^*)$ to the data set

Simulation. Prostate cancer with radiotherapy treatment

⑩ Data (synthetic data)

Prostate cancer patient treated with radiotherapy

Ctrl 1:1 Res, Day 19



* Note. We used synthetic data - hybrid Cellular Automata (CA) that track tumor growth in space with cell cycle and oxygen

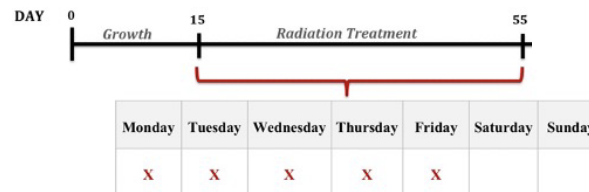
- Model $\{Y(t_i, \theta)\}$
- Consider a dynamical system that tracks tumor volume $Y(t)$ in time

$$\frac{dY}{dt} = \lambda Y \left(1 - \frac{Y}{K}\right) - \underbrace{(1 - e^{-\alpha d - \beta d^2})u(t)Y}_{\text{cell kill due to radiotherapy}},$$

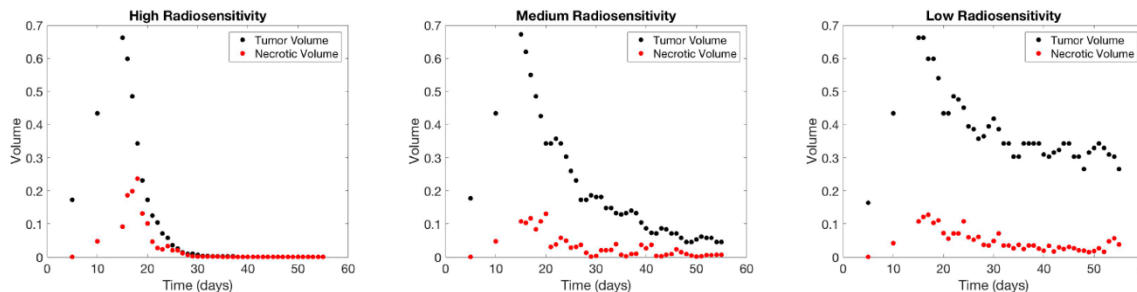
$$\theta = [\lambda, K, \alpha, \beta]$$

α, β : radiosensitivity
 λ, K : growth and capacity

- Radiotherapy schedule



- Example patients depending on response levels:



Simple scenario: One scan per week

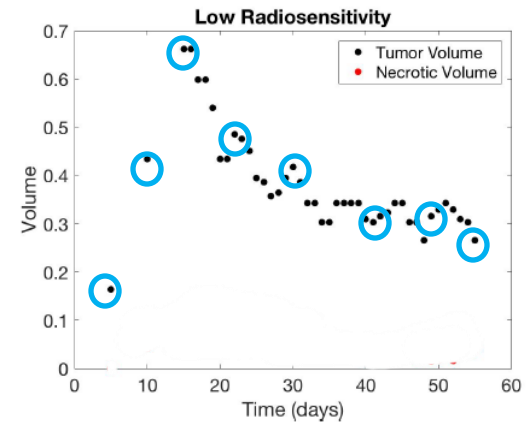
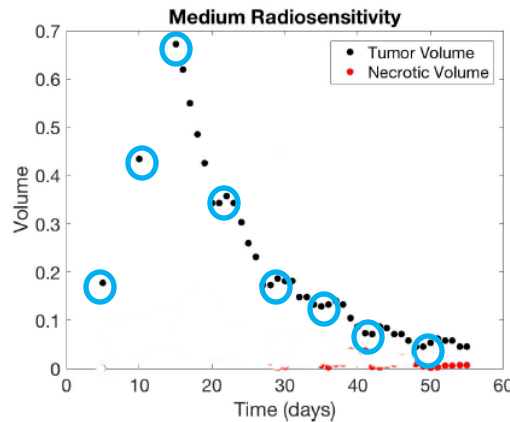
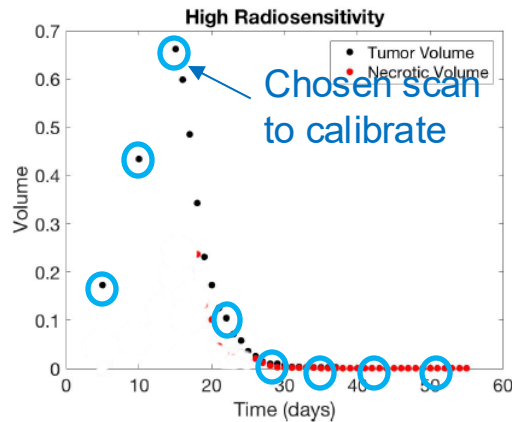
- Assume a budget of one scan per week during treatment

Use MI to choose the most informative scanning day each week

: [Mon Tue Wed Thurs Fri Sat Sun]

- Test on three different types of patients responding to radiotherapy

: High, Medium, Low responder

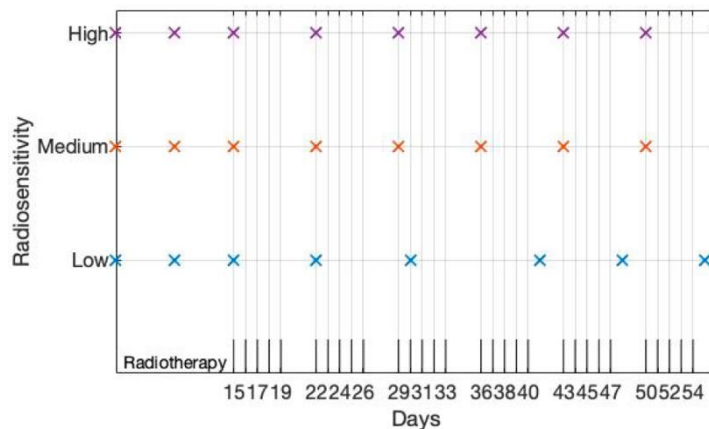


Simple scenario: Determine one scan per week

- Assume a budget of one scan per week during treatment.

Use MI to choose the most informative scanning day each week

Figure. Selected scans



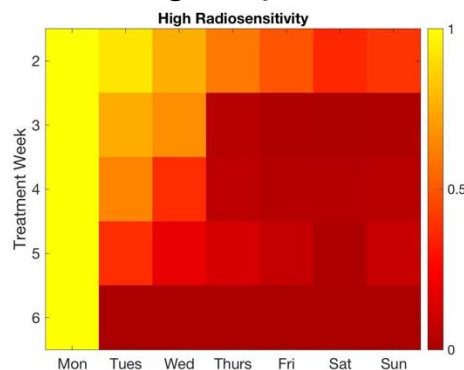
Result

First week, scan on the first day,

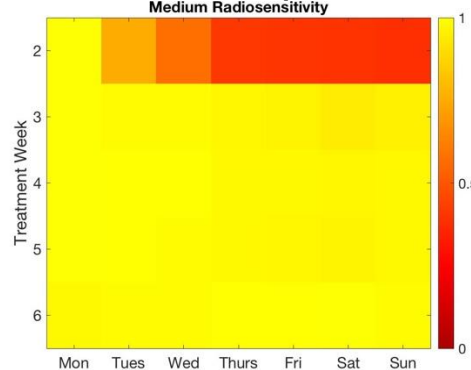
Later weeks,

scan on the first day if patient is responsive,
scan any day if patient is not responsive

Strong responder



Medium responder



Weak responder

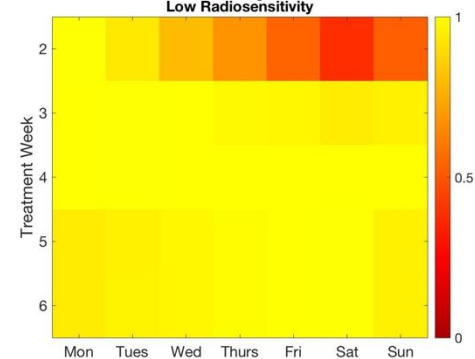


Figure. Mutual Information

PART 2. Mutual Information with temporal penalty

Application to Prostate Cancer

- 1. Synthetic data of Radiotherapy**
- 2. Clinical data of Androgen suppression treatment**

Mutual information framework for temporal data

⑩ General framework: available designs reduce by one



ex) Spatial sensor

↖ Any spatial data can be chosen

⑩ Temporal framework: lose prior designs

Day 1 ○○○○○○○○○○○○○○○ Day 14



↖ When day 5 is chosen, day 1-4 are gone.

→ Strategy: Develop an object function that rewards large mutual information while penalizing choice of scan at later times.

Modified Mutual information with temporal penalty

⑩ Objection (score) function with temporal penalty

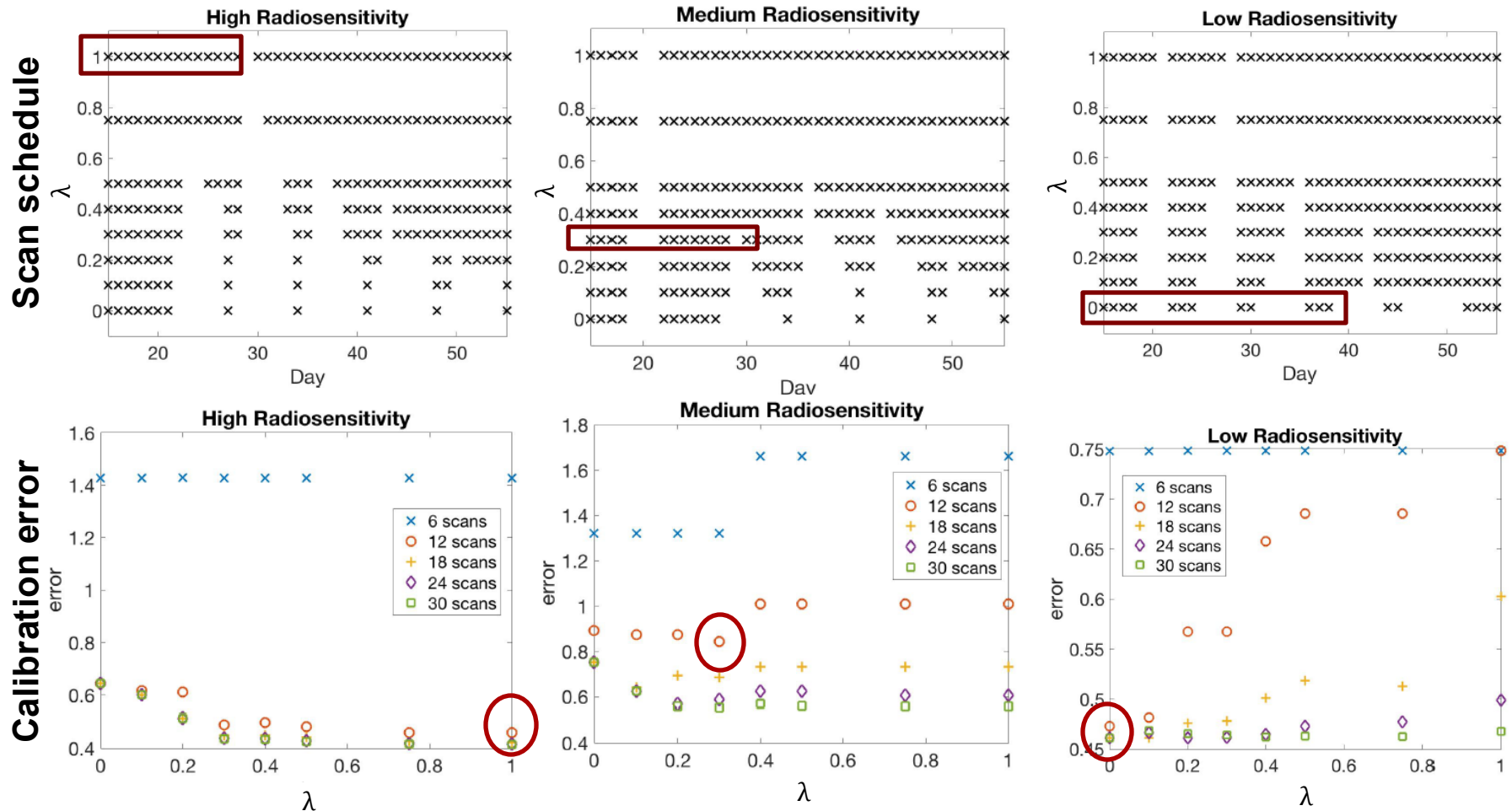
$$S_{\lambda}(i, r) = \underbrace{R(i, r)}_{\text{Rescaled MI}} - \lambda \underbrace{\frac{\sum_{j=r+1}^{i-1} R(j, r)}{\sum_{\ell=r+1}^N R(\ell, r)}}_{\text{Information Loss Ratio}},$$

⑩ Object function penalty weighted with relative tumor volume difference

$$S(i, r) = \underbrace{R(i, r)}_{\text{Rescaled MI}} - \underbrace{\frac{|\tilde{d}_r - d_N|}{\tilde{d}_r + d_N}}_{\text{Penalty Coefficient}} \cdot \underbrace{\frac{\sum_{j=r+1}^{i-1} R(j, r)}{\sum_{\ell=r+1}^N R(\ell, r)}}_{\text{Information Loss Ratio}}, \quad (2)$$

d_r : previously chosen high-fidelity measurement,
 d_N : expected final prediction using low-fidelity model

Optimized scanning schedule for three scenarios



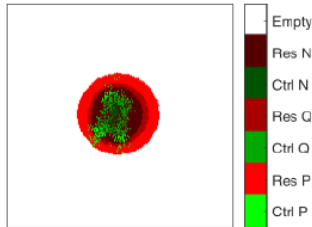
- ➔ When scan budget is low as <10 scan (high as >15 scans), $\lambda=0$ ($\lambda>0.3$) will give optimal scan schedule.
- ➔ When 12 scan budget, start with $\lambda = 0.2 \sim 0.3$, then increase (decrease) λ if patient is (not) responsive.

Simulation. Prostate cancer with radiotherapy treatment

⑩ Data (synthetic data)

Prostate cancer patient treated with radiotherapy

Ctrl 1:1 Res, Day 19



* Note. We used synthetic data - hybrid Cellular Automata (CA) that track tumor growth in space with cell cycle and oxygen

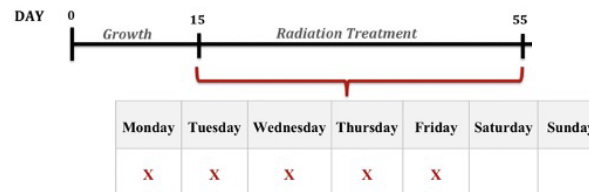
- Model $\{Y(t_i, \theta)\}$
- Consider a dynamical system that tracks tumor volume $Y(t)$ in time

$$\frac{dY}{dt} = \lambda Y \left(1 - \frac{Y}{K} \right) - \underbrace{(1 - e^{-\alpha d - \beta d^2}) u(t) Y}_{\text{cell kill due to radiotherapy}},$$

$$\theta = [\lambda, K, \alpha, \beta]$$

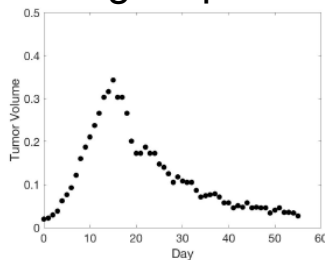
α, β : radiosensitivity
 λ, K : growth and capacity

- Radiotherapy schedule

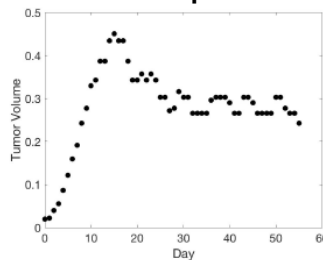


- Virtual patients depending on response levels:

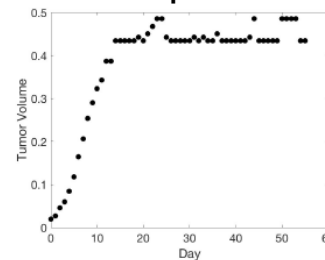
strong responder



weak responder



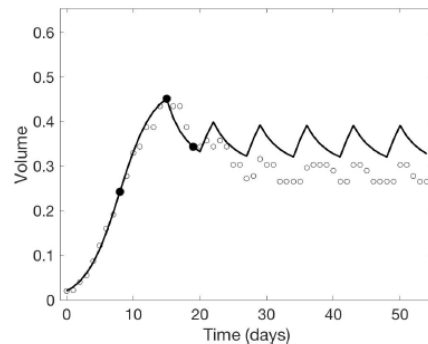
non-responder



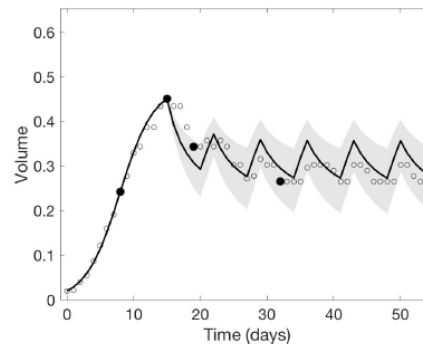
Simulation 1. Prostate cancer with radiotherapy treatment

⑩ Sequential calibration with chosen scans - weak responder.

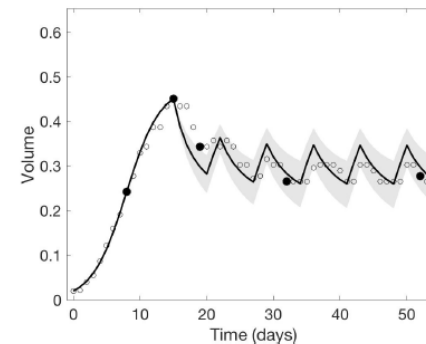
Iteration 1



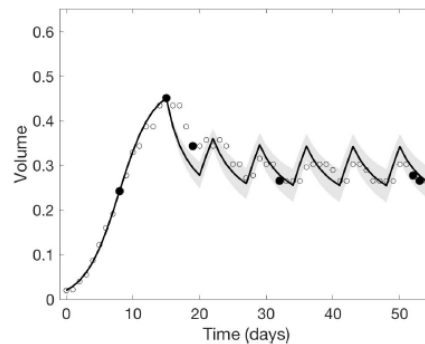
Iteration 2



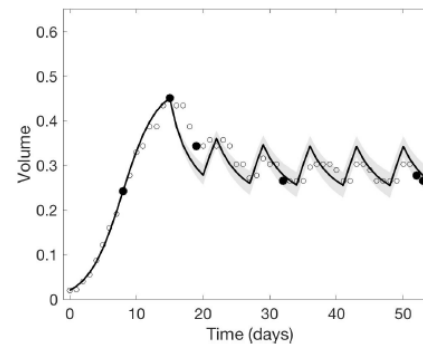
Iteration 3



Iteration 4



Iteration 5



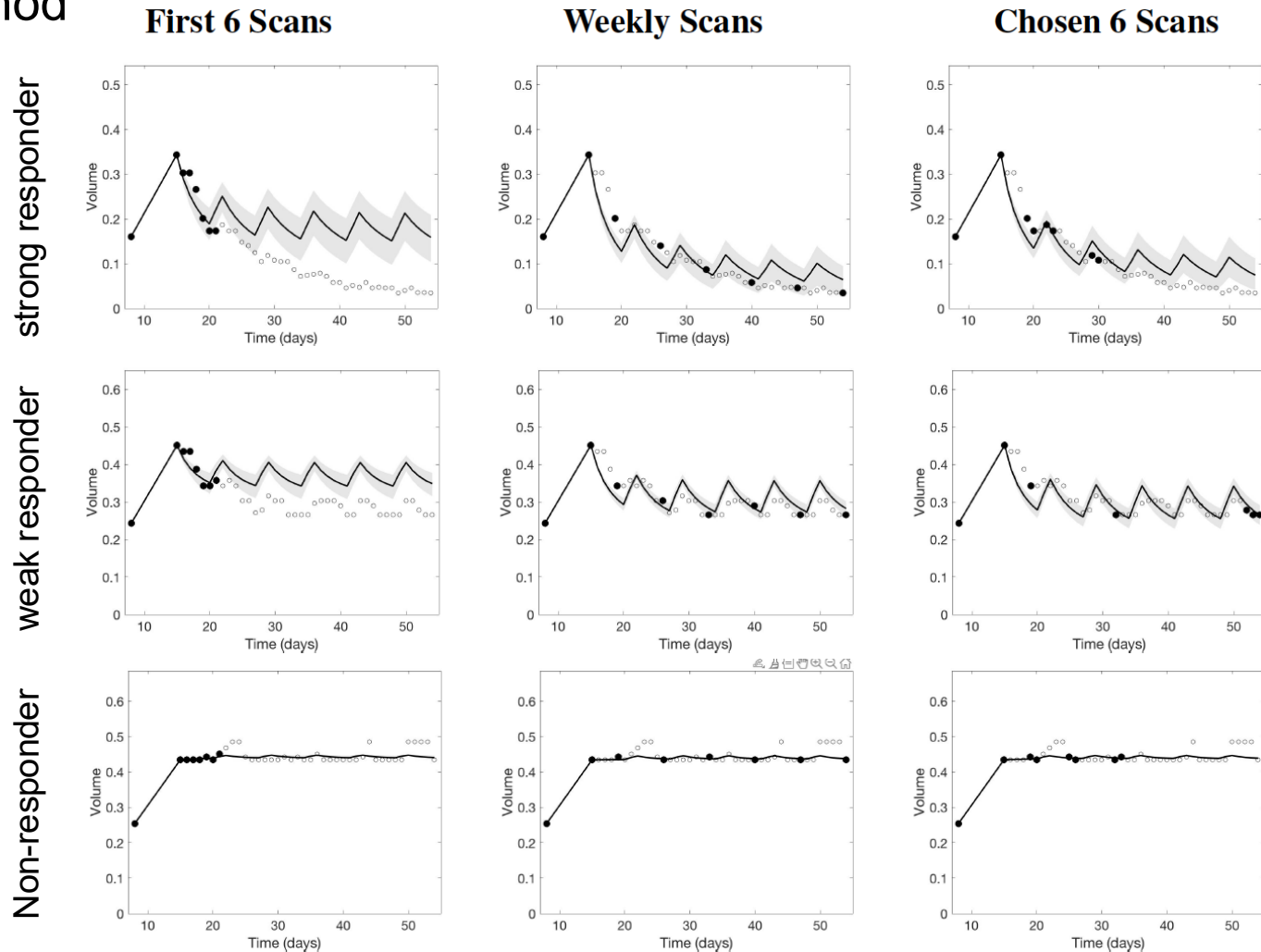
*Black dots are
the chosen scans

→ Let us compare the following with 6 (+1-2 for initial growth) scans budget :

1. First 6 scans (Mon-Sat on week 1)
2. Weekly scans (Fridays for 6 weeks)
3. Chosen schedule of **6 scans** with our framework

Simulation 1. Prostate cancer with radiotherapy treatment

⑩ Comparison between 1) First 6 scans, 2) Weekly scans, 3) our method



Simulation 1. Prostate cancer with radiotherapy treatment

⑩ Comparison between 1) First 6 scans, 2) Weekly scans, 3) our method

| | | First Scans | Weekly Scans | Chosen Scans |
|--------|--------------------|-------------|--------------|--------------|
| Strong | Error | 0.0102 | 0.0013 | 0.0017 |
| | Uncertainty | 3.5445 | 2.4024 | 2.4587 |
| Weak | Error | 0.0056 | 0.0012 | 0.0010 |
| | Uncertainty | 2.0207 | 1.2887 | 1.7847 |
| Non | Error | 0.0003 | 0.0003 | 0.0003 |
| | Uncertainty | 0.3412 | 0.2649 | 0.2409 |

$$\text{Error : mean square error : } \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i; \theta))^2$$

Uncertainty : area of 95% credible intervals

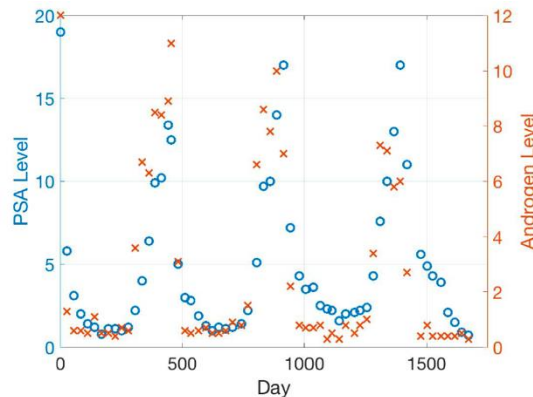
→ Our chosen scan schedule is doing comparable to the equally spaced scan schedule in terms of the accuracy and uncertainty,

Simulation 2. Prostate cancer with androgen treatment

⑩ Data (clinical data)

Prostate cancer patient treated with intermittent androgen suppression therapy

- Prostate-Specific Antigen (PSA) level
- Serum Androgen level



Data from N. Bruchovsky et al (2006)

⑩ Model

$$\frac{dx_1}{dt} = \max\left(\mu\left(1 - \frac{q_1}{Q}\right)x_1, 0\right) - dx_1(x_1 + x_2) - c\frac{K}{Q + K}x_1$$

$$\frac{dx_2}{dt} = \max\left(\mu\left(1 - \frac{q_2}{Q}\right)x_2, 0\right) - dx_2(x_1 + x_2) + c\frac{K}{Q + K}x_1$$

$$\frac{dQ}{dt} = m(A - Q) - \frac{\mu(Q - q_1)x_1 + \mu(Q - q_2)x_2}{x_1 + x_2}$$

$$\frac{dA}{dt} = \gamma_1 u(t) \left(1 - \frac{A}{A_0}\right) + \gamma_2 - \delta A$$

$$\frac{dP}{dt} = bQ + \max\left(\sigma_1\left(1 - \frac{q_1}{Q}\right)x_1, 0\right) + \max\left(\sigma_2\left(1 - \frac{q_2}{Q}\right)x_2, 0\right) - \epsilon P,$$

- $x_1(t)$, $x_2(t)$: prostate cancer population
- $P(t)$: Prostate-Specific Antigen(PSA) level
- $A(t)$: Serum Androgen level
- $Q(t)$: Intermediate Androgen level

[Dr. Kuang group] W Meade, et al. (2022)

⑩ Collect data for 1.5 cycle to predict 3.5 cycle

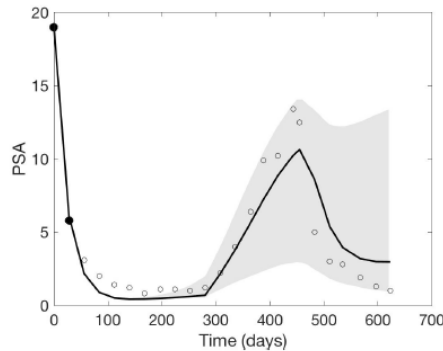
⑩ Test cases :

1. Collect PSA data only for 1.5 cycle
2. Collect both PSA and Androgen data for 1.5 cycle

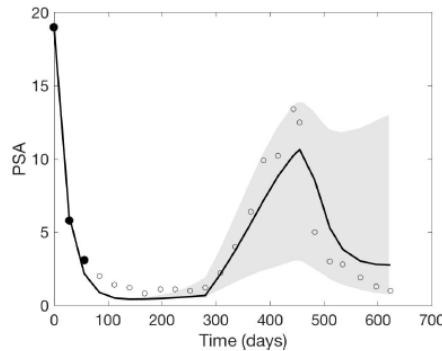
Simulation 2. Prostate cancer with androgen treatment

⑩ Sequential calibration procedure of only PSA data for 1.5 cycle

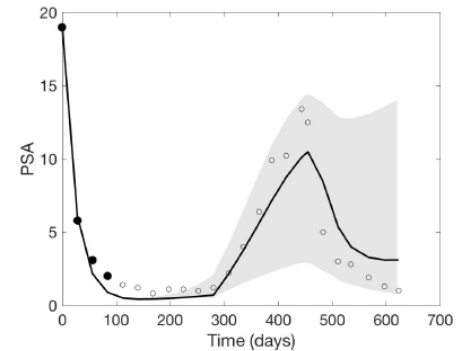
Iteration 1



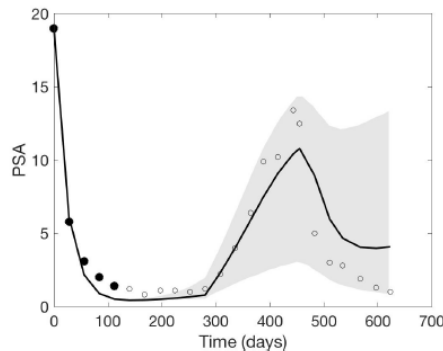
Iteration 2



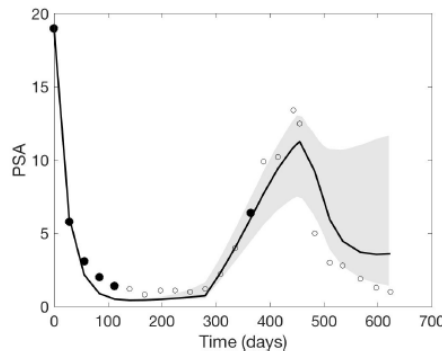
Iteration 3



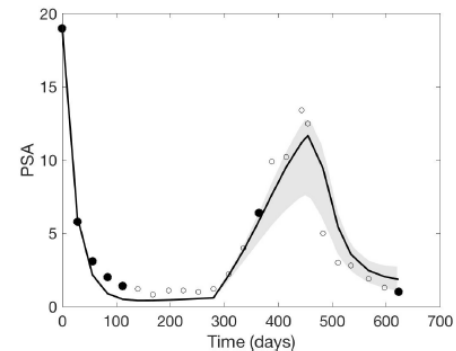
Iteration 4



Iteration 5



Iteration 6



→ Let us compare the following with 5 (+2 for initial calibration) scans budget :

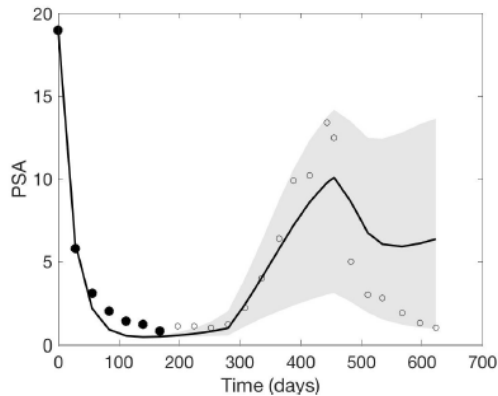
1. First 5 scans (every 28 days)
2. Evenly-spaced scans (every 140 days)
3. Chosen schedule of **5 scans** with our framework

*Black dots are
the chosen scans

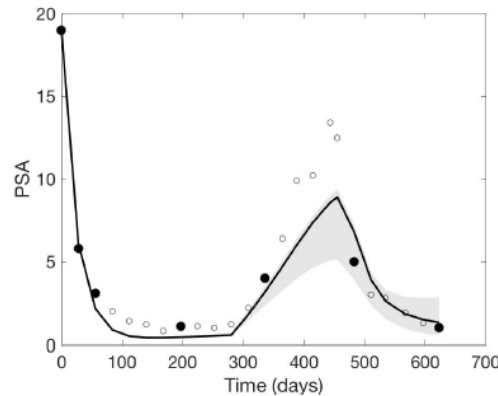
Simulation 2. Prostate cancer with androgen treatment

⑩ Comparison of PSA data only collection for 1.5 cycle

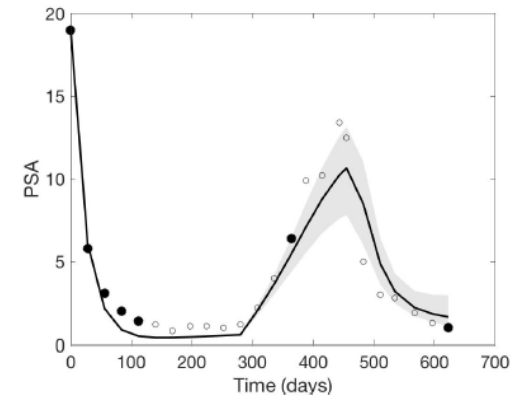
First 5 Scans



Evenly-Spaced 5 Scans



Chosen 5 Scans

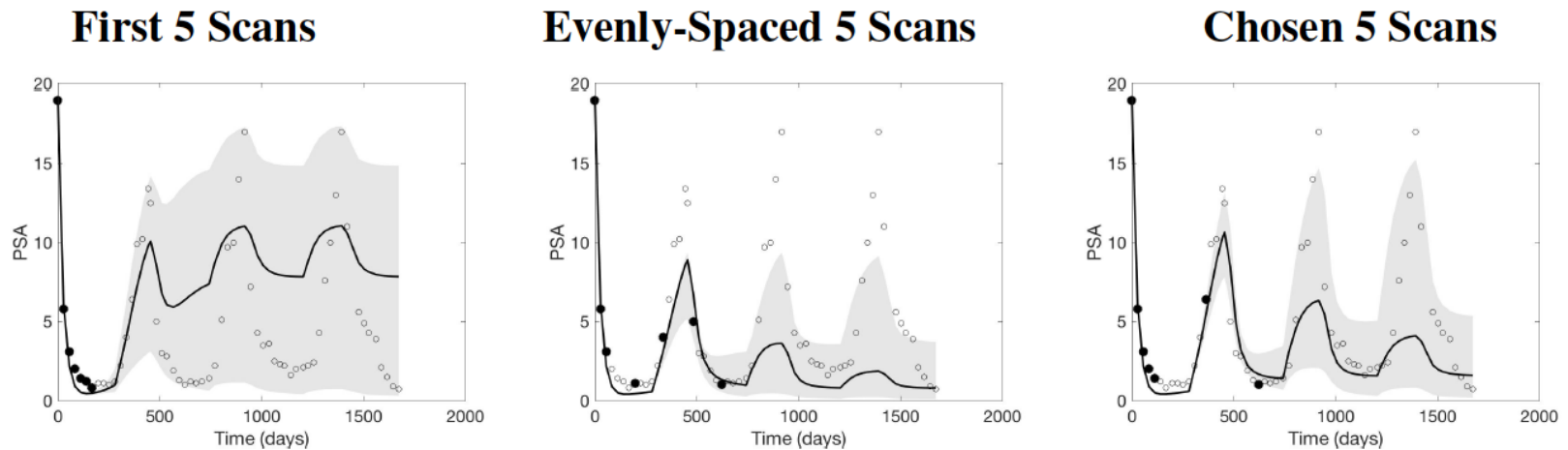


| | First Scans | Evenly-Spaced Scans | Chosen Scans |
|--------------------|--------------------|----------------------------|---------------------|
| Error | 5.80 | 3.04 | 1.92 |
| Uncertainty | 126.54 | 31.18 | 36.13 |

→ Our adaptive scan schedule gives most accurate result.

Simulation 2. Prostate cancer with androgen treatment

⑩ Comparison of PSA data only collection, prediction up to 3.5 cycle



| | First Scans | Evenly-Spaced Scans | Chosen Scans |
|-------------|-------------|---------------------|--------------|
| Error | 17.98 | 18.41 | 10.85 |
| Uncertainty | 665.15 | 212.81 | 297.95 |

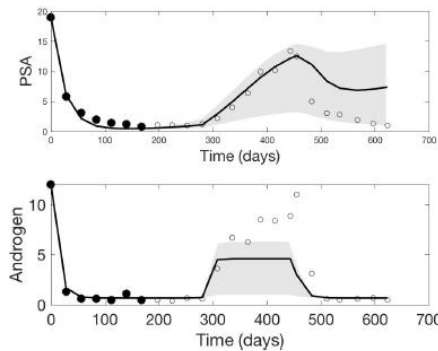
→ Our adaptive scan schedule gives most accurate result.

Simulation 2. Prostate cancer with androgen treatment

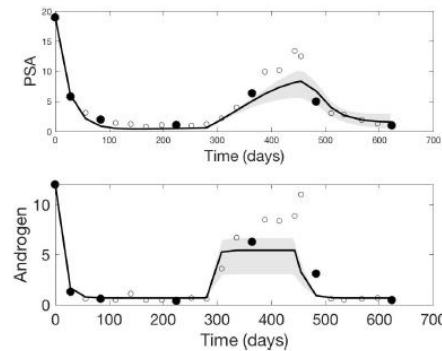
⑩ Comparison of PSA & Androgen data collection for 1.5 cycle

| | 0 | 28 | 56 | 84 | 112 | 140 | 168 | 197 | 224 | 252 | 280 | 308 | 336 | 364 | 388 | 415 | 443 | 455 | 483 | 511 | 535 | 567 | 597 | 623 |
|----------|---|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| PSA | X | X | X | X | | | | | | | | | | | | X | X | X | X | | | | | |
| Androgen | X | X | | X | | | | | | | | X | | | | X | X | | | | | | | |

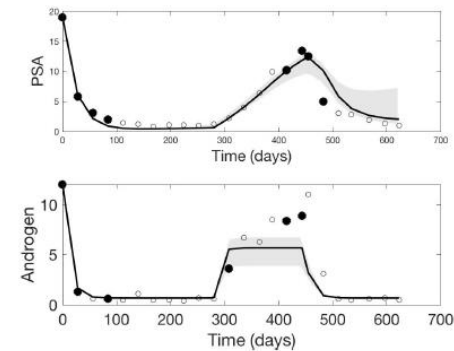
First 10 Scans



Evenly-Spaced 10 Scans



Chosen 10 Scans

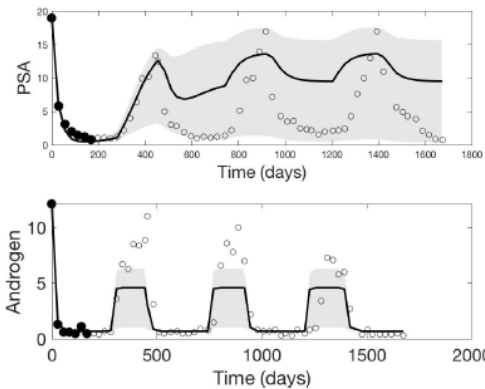


| | First Scans | Evenly-Spaced Scans | Chosen Scans |
|-------------|-------------|---------------------|--------------|
| Error | 13.29 | 7.42 | 6.05 |
| Uncertainty | 4231.66 | 1413.62 | 1575.57 |

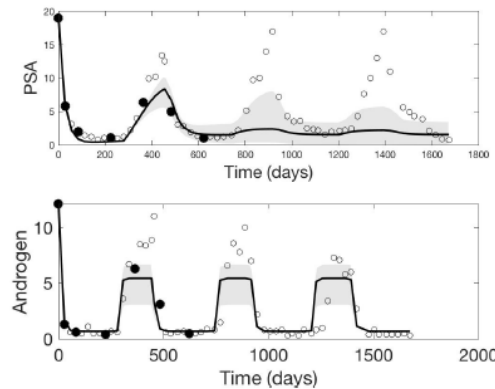
Simulation 2. Prostate cancer with androgen treatment

⑩ PSA & Androgen data collection, prediction up to 3.5 cycle

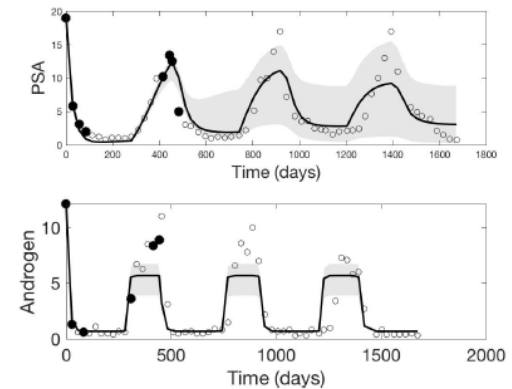
First 10 Scans



Evenly-Spaced 10 Scans



Chosen 10 Scans



| | First Scans | Evenly-Spaced Scans | Chosen Scans |
|--------------------|--------------------|----------------------------|---------------------|
| Error | 33.47 | 22.45 | 7.62 |
| Uncertainty | 22510.54 | 7223.40 | 12741.40 |

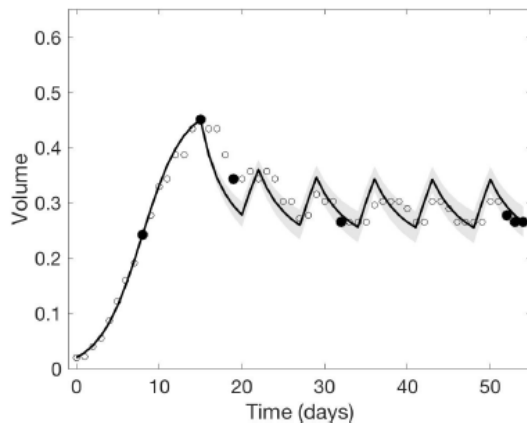
→ Our adaptive scan schedule gives most accurate result.

PART 3. Further ideas to reduce # of scans

Application to Prostate Cancer with Radiotherapy

Simulation. Prostate cancer with radiotherapy treatment

⑩ Example of selected scans in weak responder case



- Model $\{Y(t_i, \theta)\}$
- Consider a dynamical system that tracks tumor volume $Y(t)$ in time

$$\frac{dY}{dt} = \lambda Y \left(1 - \frac{Y}{K}\right) - \underbrace{(1 - e^{-\alpha d - \beta d^2})u(t)Y}_{\text{cell kill due to radiotherapy}},$$

$$\theta = [\lambda, K, \alpha, \beta]$$

α, β : radiosensitivity
 λ, K : growth and capacity

3 scans required for initial estimate of parameters

→ Use parameter prior distribution estimated from population

Last scans are redundant

→ Check convergence criteria of parameter value and posterior

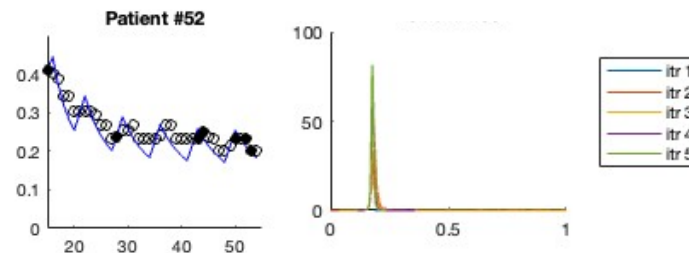
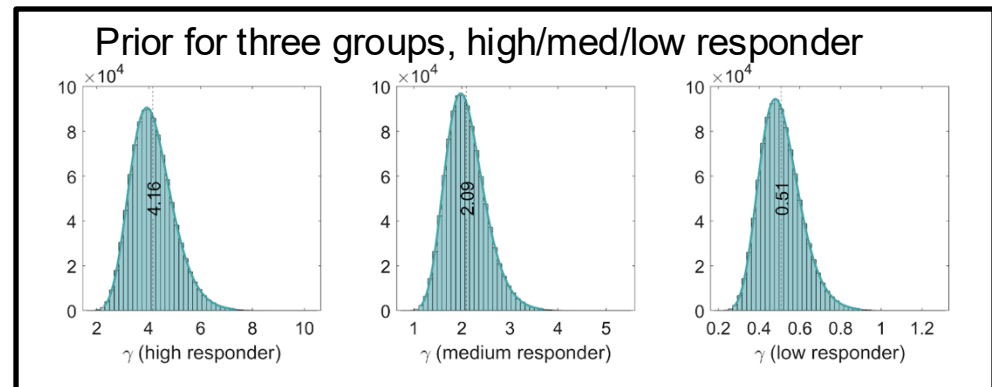
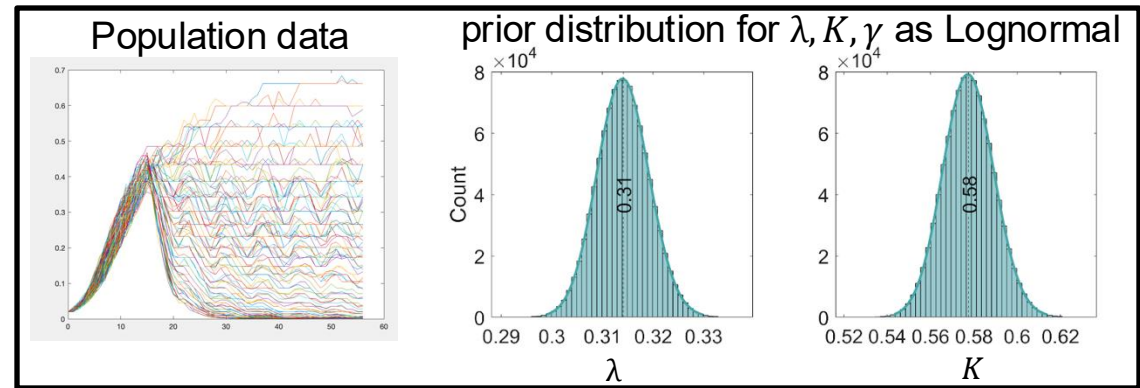
Workflow of patient scanning protocols

Pre-estimate population-level prior distribution

Select next scan using one data point

Based on the next scan, update prior among groups

Continue to select scans until Max budget OR convergence of posterior distribution

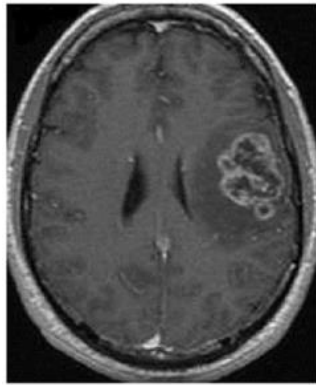


- Prior distribution computed by Monolix (nonlinear mixed-effects model)

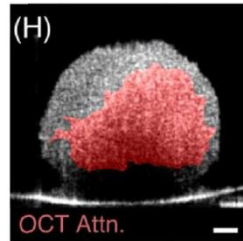
More data - Adding model complexity

- Imaging necrotic region

: data of total tumor and necrotic volume →



MRI, L. Han et al, MBE (2019)



Y. Huang et al, Optical Coherence Tomography Detects Necrotic Regions and Volumetrically Quantifies Multicellular Tumor Spheroids, *Cancer Res* (2017)

- Model :

Dynamical system tracking tumor (V) and necrotic (N) volume

$$\frac{dV}{dt} = \lambda V \left(1 - \frac{V}{K}\right) - \eta V$$

$$\frac{dN}{dt} = \eta V - \zeta N$$

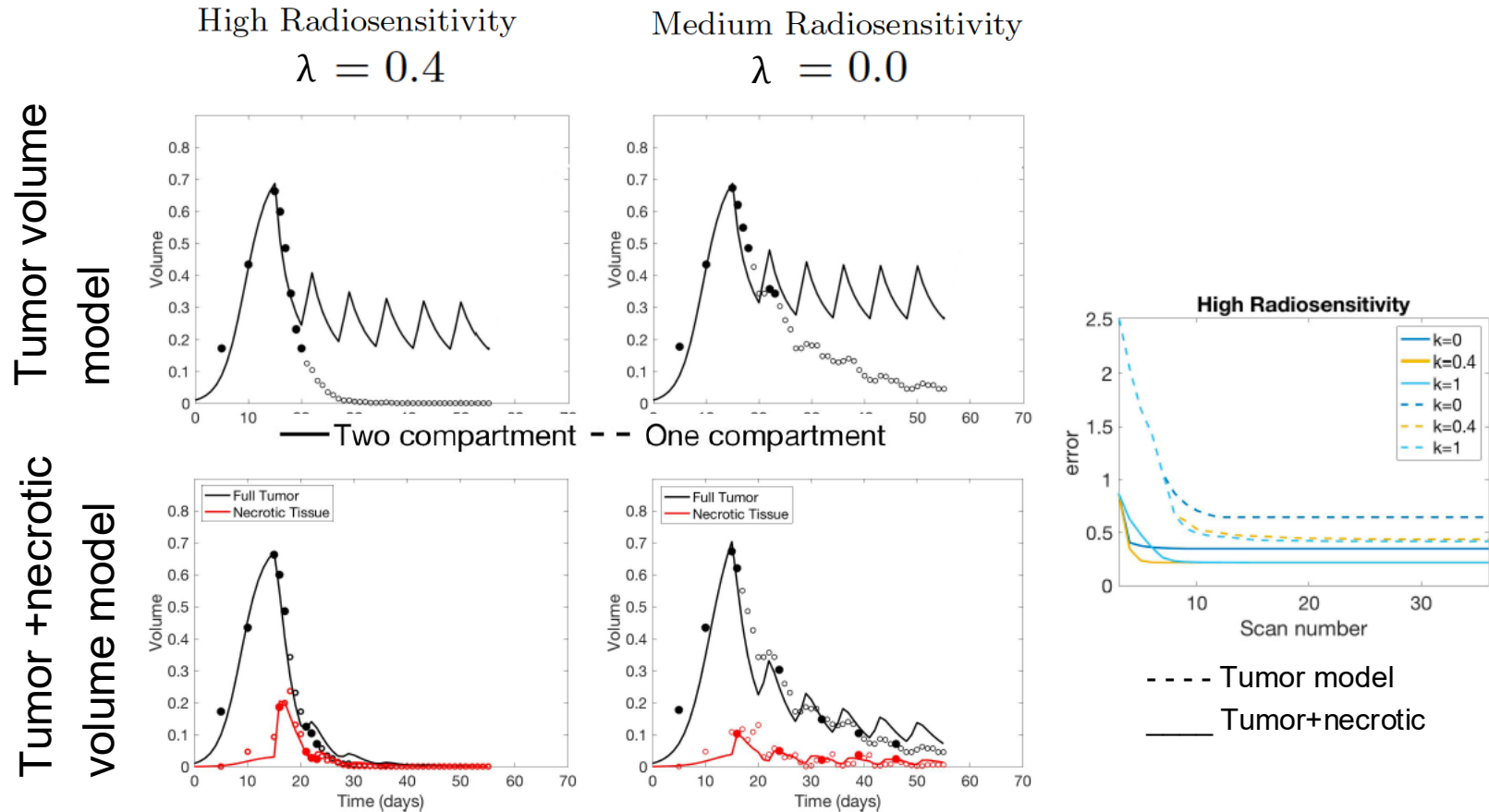
- Reference demonstrating that tracking dead matter is important in radiotherapy

T.D. Lewin, H.M. Byrne et al. The importance of dead material within a tumour on the dynamics in response to radiotherapy. *Physics in Medicine and Biology* (2019).

T.D. Lewin, P.K. Maini et al. A three-phase model to investigate the effects of dead material on the growth of avascular tumours. *Mathematical Modelling of Natural Phenomena* (2019).

Improved prediction error with necrotic volume scan

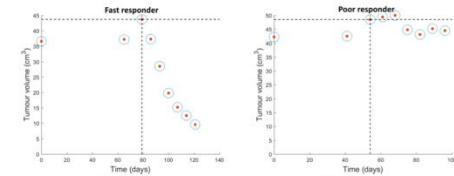
- With fixed 6 scan budget (+ 2 for initial growth)



➔ Measuring necrotic volume in addition to tumor volume, and proposed scan schedule gives much accurate calibration only with 6 scans after treatment.

Current/Future work

- Validation to more clinical data!
- Incorporate multiple collection modes of data
 - MRI imaging, Cerebrospinal fluid (CSF) collection, Bone marrow biopsy, blood draw
- Incorporate multiple models with different fidelity levels
 - Hierarchy of ODE and PDE models
- Develop efficient and accurate theoretical and computational approaches for conditional mutual information.
- Other Bayesian optimization methods for experimental design



Radiotherapy Patient data (Lewin et al., 2022)

- H Cho, A. Lewis, K Storey (2020)) *J. Clin. Med.* 9, 3208
- H Cho, A. Lewis, K Storey, et al. (2023) *Math. Bio. Eng.*, 20(10)

Joint work with

Russ Rockne (Beckman Research Institute, CoH)

Ya-Huei Kuo (Beckman Research Institute, CoH)

Margarita Gutova (Beckman Research Institute, CoH)

Vikram Adhikarla (Beckman Research Institute, CoH)

Michael Barish (Beckman Research Institute, CoH)



Sihem Cheloufi (UC Riverside)

Weitao Chen (UC Riverside)

Qixuan Wang (UC Riverside)

Students at UC Riverside

Tony Li

Austin Hansen

Kevin Su

Khoi Vo

Allison Lewis (Lafayette College)

Katie Storey (Lafayette College)

Tin Phan (LANL)

Funding

SIMONS
FOUNDATION



UC RIVERSIDE

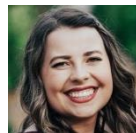
Thank you

Thanks to



Students at UC Riverside

Ph.D.



Tony Li Austin Hansen Daley Thomale



Dr. Russell
Rockne



Dr. Margarita Gutova



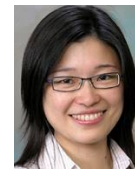
Dr. Vikram Adhikarla



Dr. Michael
Barish.



Dr. Christian
Brown



Dr. Ya-Huei Kuo



Kevin Su Minh Vu



Dr. Allison Lewis
Lafayette Univ



Dr. Kathleen Storey
Lafayette Univ.



Dr. Sihem Cheloufi
UC Riverside



Dr. Sandhya Prabhakaran
Moffitt Cancer Centre



Funding



Application/Simulation to Prostate Cancer

1. Synthetic data of Radiotherapy
2. Clinical data of Androgen suppression treatment

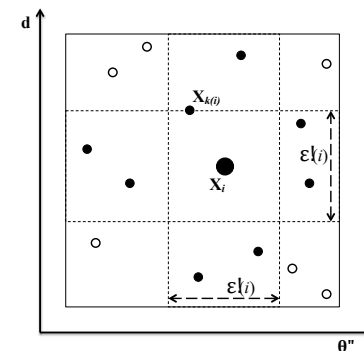
kNN estimate of Mutual Information

$$I(\theta; d_n | D_{n-1}, \xi_n) = \int_{\mathcal{D}} \int_{\Omega} p(\theta, d_n | D_{n-1}, \xi_n) \log \frac{p(\theta, d_n | D_{n-1}, \xi_n)}{p(\theta | D_{n-1}) p(d_n | D_{n-1}, \xi_n)} d\theta dd_n$$

Methods for approximating MI:

- Monte Carlo sampling - prohibitively expensive for moderate to high-dimensional problems!
- k^{th} -Nearest Neighbor (kNN) estimate

Kraskov et al. (2004) Estimating mutual information, Phys. Rev. E



$$I(\theta; d_n | D_{n-1}, \xi_n) \approx \psi(k) - \frac{1}{N} \left[\sum_{i=1}^N \psi(n_{\theta}(i) + 1) + \sum_{i=1}^N \psi(n_d(i) + 1) \right] + \psi(N)$$

Delayed Rejection Adaptive Metropolis Algorithm

1. Determine $\theta^0 = \arg \min_{\theta} \sum_{i=1}^N [y_i - f(t_i, \theta)]^2$.

2. Construct covariance estimate V .

3. For $k = 1, \dots, M$

(a) Construct candidate $\theta^* \sim N(\theta^{k-1}, V)$.

(b) Compute

$$SS_{\theta^*} = \sum_{i=1}^N [y_i - f(t_i, \theta^*)]^2,$$

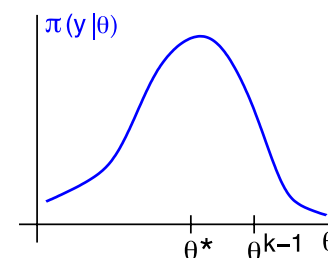
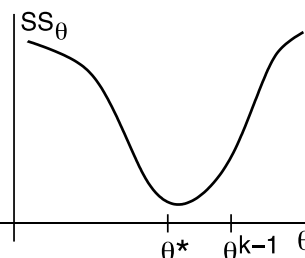
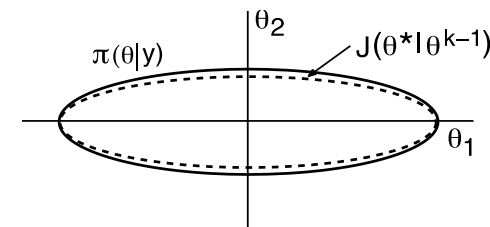
$$\pi(y|\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-SS_{\theta}/2\sigma^2}.$$

(c) Compute $\alpha(\theta^*|\theta^{k-1}) = \min(1, e^{-[SS_{\theta^*} - SS_{\theta^{k-1}}]/2\sigma^2})$.

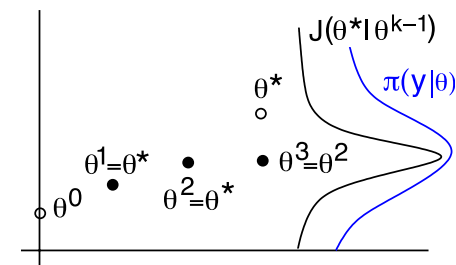
(d) Accept θ^* with probability α .

Proposal Distribution

$$J(\theta^*|\theta^{k-1}) = N(\theta^{k-1}, V)$$



Note: Minimizing the SSQ function is equivalent to maximizing the likelihood



H. Haario et al. (2006) DRAM: Efficient adaptive MCMC, Stats. Comput.