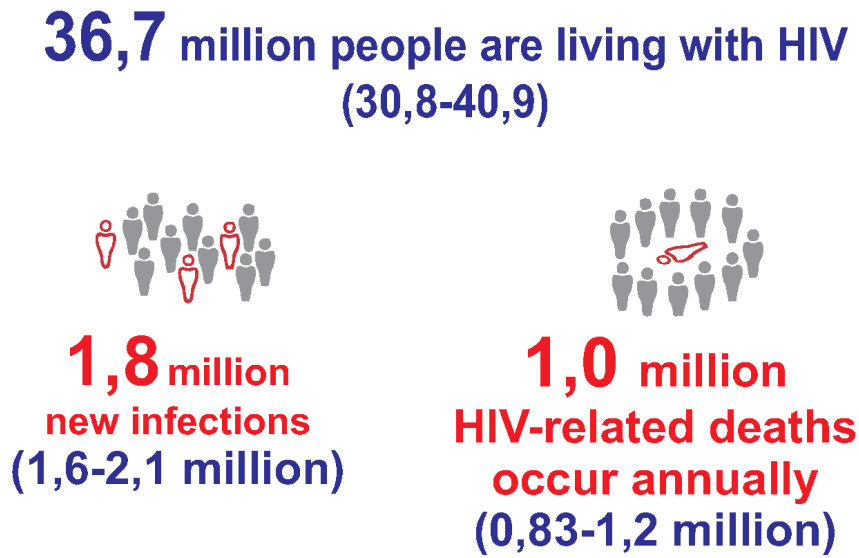
Machine learning algorithms for inverse problems
Time-adaptive parameter identification in mathematical model of HIV infection with drug therapy
Lecture 10

Outline

- **Model problem** System of ODE describing dynamics of HIV infection under treatment of a reverse transcriptase inhibitor [SBC].
- **Inverse problem:** Find the efficacy of the drug, given time-dependent noisy observations of the solution of the model problem. [BG]
- **Aim:** comparison of different techniques (analytic approach, time-adaptive fem [BG], least squares approach and machine learning methods) for efficient reconstruction of the efficacy of the drug in the model problem.

[ME] M. Eriksson, Parameter identification in a mathematical model of HIV infection with drug therapy, Master thesis, https://gupea.ub.gu.se/handle/2077/54664, 2017.
[SBC] P.K. Srivastava, M. Banerjee and P. Chandra, Modeling the drug therapy for HIV infection, *Journal of Biological Systems*, **17**, 213-223, 2009.
[BG] L. Beilina, I. Gainova, *Time-adaptive FEM for distributed parameter identification in mathematical model of HIV infection with drug therapy*, Inverse Problems and Application, Springer Proceedings in Mathematics and Statistics, vol. 120, Springer, 2015, pp. 111-124.
[BG2] L. Beilina, I. Gainova, *Time-adaptive optimization in a parameter identification problem of HIV infection*, arXiv:1912.01112v1

Global HIV epidemic (2016)



Source: UNAIDS/WHO estimates

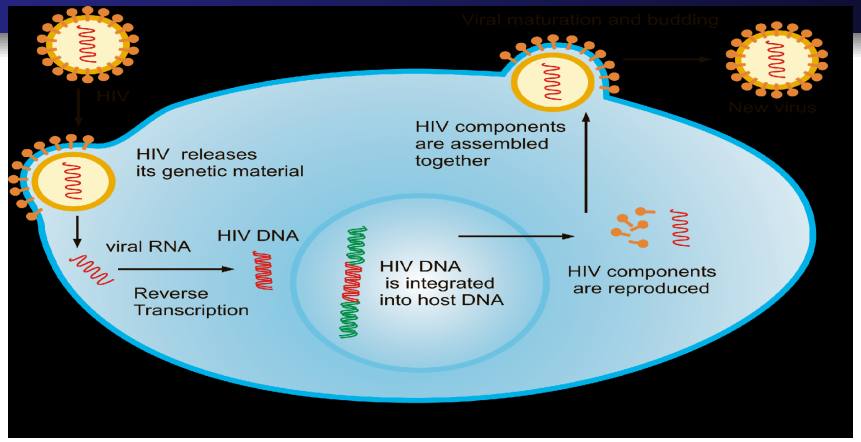
- Human Immunodeficiency Virus (HIV) is one of the most infectious and dangerous viral agents which still remains a major public health challenge in the world.
- According to UNAIDS for 2016, 36.7 million people living with HIV and about 2.1 million new infections were recorded in 2015.

UNAIDS. (2017). The Joint United Nations Programme on HIV and AIDS. Fact sheet July 2017 Global HIV statistics. Retrieved on January 20, 2018. http://www.unaids.org/sites/default/files/media asset/UNAIDS FactSheet en.pdf.

The main features of HIV

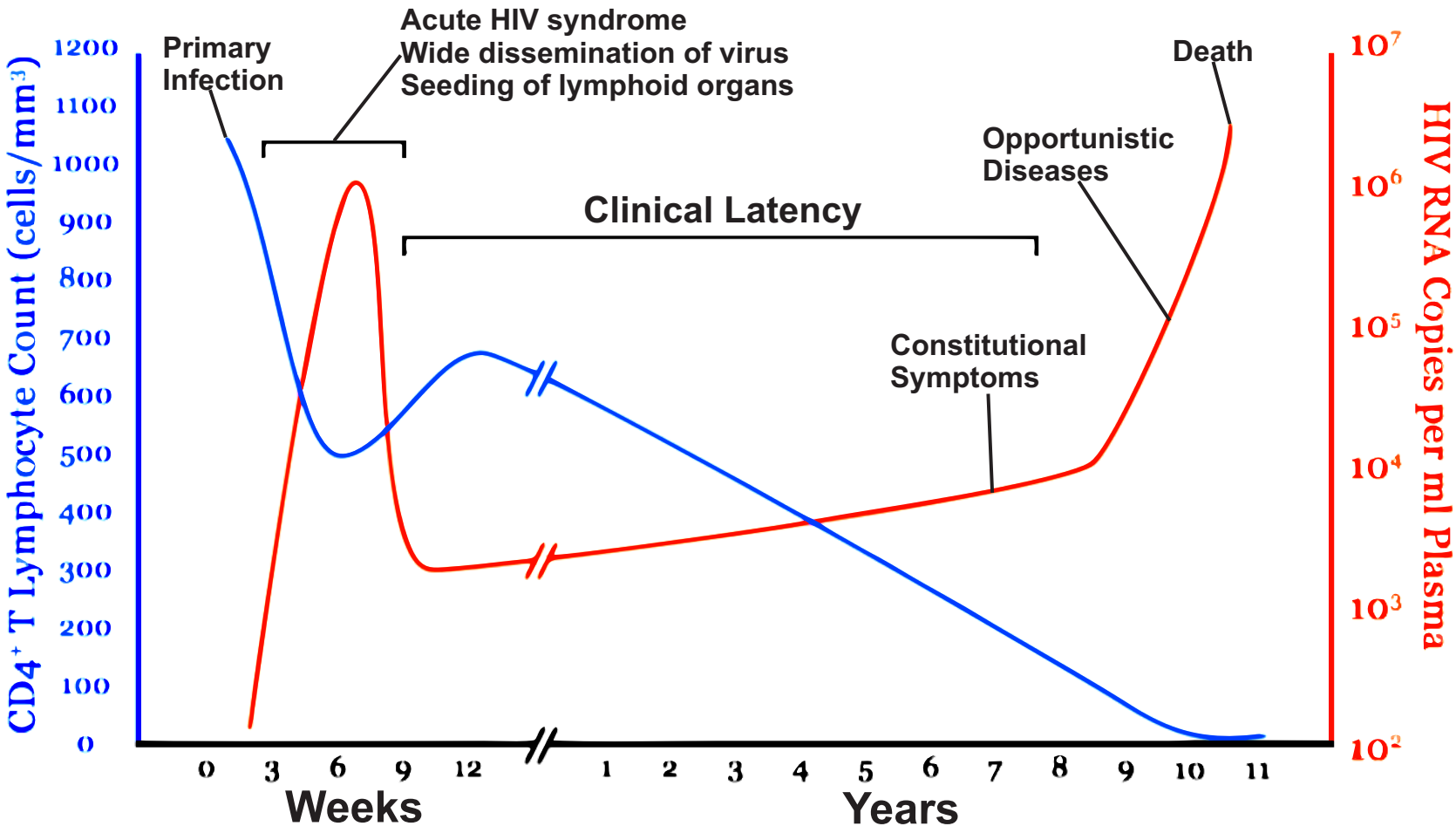
- HIV attacks the immune system cells that have CD4 receptors, such as T-lymphocyte-helper cells, macrophages, dendritic cells, etc. As a result, the immune system is depleted and the tissues of the lymphoid organs are destroyed.
- The persistence of latent (asymptomatic) infection is an important feature of the pathogenesis of HIV infection: constant presence of large reservoirs of latently infected cells is one of the main obstacles of treatment of HIV infection.
- HIV differs from other viruses by a high mutation level, the mutation rate is $10^{-5} - 10^{-4}$ per nucleotide during one replication cycle.
- This allows the virus to "escape" from humoral and cellular defense factors of our immune system and to form multiplicity of drug-resistant strains.
- Paradoxical feature of HIV is that activation of the immune system does not lead to a suppression of virus multiplication, but to opposite to activation of latently infected cells, which start to produce new viral particles.

HIV life-cycle



1. HIV first locates the CD4 cell, attaches to its surface and then releases its genetic material (viral RNA) and enzymes into the CD4 cell.
2. The enzyme reverse transcriptase (RT) copies the viral RNA into viral DNA.
3. The viral DNA is integrated in to the CD4 cellâs nuclear material.
4. The individual components of HIV are then produced within the CD4 cell.
5. The individual components of HIV are then assembled together to make new HIV viruses.
6. New viruses are released from the CD4 cell. These infect other CD4 cells where the cycle repeats itself.

Three phases of HIV infection



OâBrien, S., Hendrickson, S. Host genomic influences on HIV/AIDS, Genome Biology 14:201, (2013).

Highly active antiretroviral therapy (HAART)

- In 1996-1997 started active antiretroviral therapy (HAART) treatment of HIV which was based on the use of a combination of antiviral drugs. HAART led to a significant improvement in the quality of life of patients, has caused a clear decrease in AIDS-related diseases and mortality.
- At the present stage, the development of an optimal HAART strategy is impossible without using of methods of mathematical modelling and mathematical programming, due to a complex combinatorics of the drugs used and, as a result, the appearance of many drug-resistant strains. For example, a treatment scheme known as the mega-therapy (MDRT — multi-drug rescue therapy) may include combinations of 9 to 15 antiretroviral drugs.

The role of mathematical methods

- development of personalized therapy that takes into account individual characteristics of a patient;
- development of complex combinatorial treatment schemes, studies of the joint action of drugs;
- assessing the toxicity of applied drugs for the patient;
- development and research of new treatment strategies;
- identification of unknown parameters that cannot be measured experimentally;
- processing and analysis of large data arrays: both recorded clinical cases and experimental data obtained using modern measurement methods (BIG DATA ANALYSIS);
- assessing the cost of treatment and making recommendations at the governmental levels to prevent the spread of the disease over the world.

Model problem: The forward problem

Let $\Omega = [0, T]$ be the time domain of our problem. The model is given by the following system of ODE for $u(0) = (u_1^0, u_2^0, u_3^0, u_4^0)$:

$$\begin{cases} \dot{u}_1 = s - k u_1 u_4 - \mu u_1 + (\eta\alpha + b)u_2, \\ \dot{u}_2 = k u_1 u_4 - (\mu_1 + \alpha + b)u_2, \\ \dot{u}_3 = (1-\eta)\alpha u_2 - \delta u_3, \\ \dot{u}_4 = N\delta u_3 - c u_4, \\ u_1(0) = u_1^0 = 300\ mm^{-3}, u_2(0) = u_2^0 = 10\ mm^{-3}, \\ u_3(0) = u_3^0 = 10\ mm^{-3}, u_4(0) = u_4^0 = 10\ mm^{-3}. \end{cases} \tag{1}$$

where u_1 = healthy T cells, u_2 = pre-RT infected cells, u_3 = post-RT infected cells and u_4 = virus.
In green are biological constants, known at great accuracy, see Table. In black is the solution of the model problem. The function u_4 is known at a subset of the time domain. In red is the unknown function $\eta(t)$.

[SBC] P.K. Srivastava, M. Banerjee and P. Chandra, Modeling the drug therapy for HIV infection, *Journal of Biological Systems*, **17**, 213-223, 2009.

Known parameters in the model problem

Parameter	Value	Units	Description
s	10	$mm^{-3}day^{-1}$	inflow rate of T cells
μ	0.01	day^{-1}	natural death rate of T cells
k	2.4E-5	$mm^{3}day^{-1}$	interaction-infection rate of T cells
μ_1	0.015	day^{-1}	death rate of infected cells
α	0.4	day^{-1}	transition rate from pre-RT to post-RT cla
b	0.05	day^{-1}	reverting rate of infected cells to uninfecte
δ	0.26	day^{-1}	death rate of actively infected cells
c	2.4	day^{-1}	clearance rate of virus
N	1000	$vir/cell$	total number of viral particles

Model problem: The forward problem

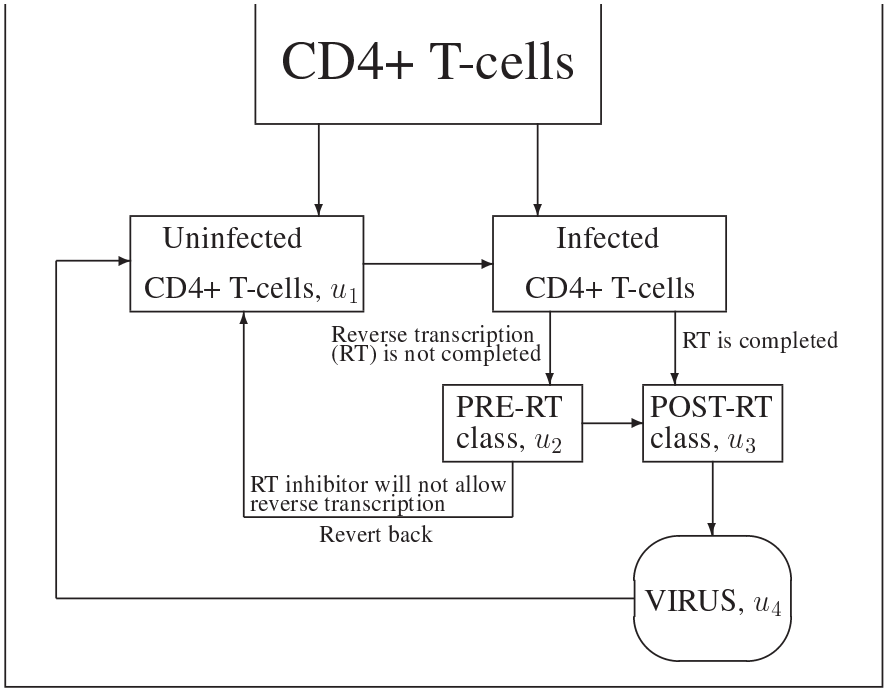


Figure 1: *The effect of Reverse Transcriptase Inhibitor (RTI) on the dynamics of HIV infection.*

$$\begin{cases} \dot{u}_1 = s - k u_1 u_4 - \mu u_1 + (\eta\alpha + b)u_2, \\ \dot{u}_2 = k u_1 u_4 - (\mu_1 + \alpha + b)u_2, \\ \dot{u}_3 = (1-\eta)\alpha u_2 - \delta u_3, \\ \dot{u}_4 = N\delta u_3 - c u_4. \end{cases} \tag{2}$$

Parameter Identification Problem

Let $M_\eta = \{\eta \in C(\Omega_T) : \eta(t) \in (0,1)\ \forall t \in \Omega_T\}$ be the set of admissible functions for η. The inverse, or parameter identification problem, is defined as follows:

Parameter Identification Problem (PIP)
Determine $\eta(t) \in M_\eta$, assuming the following function is known

$$u_4(t) = g(t), t \in \Omega_{obs} \subseteq \Omega_T \tag{3}$$

The function $g(t)$ represents (noisy) observations of the virus function inside the set Ω_{obs}.

[BG] L. Beilina, I. Gainova, *Time-adaptive optimization in a parameter identification problem of HIV infection*, arXiv:1912.01112v1

Analytic reconstruction of η

Recall the model ODE system:

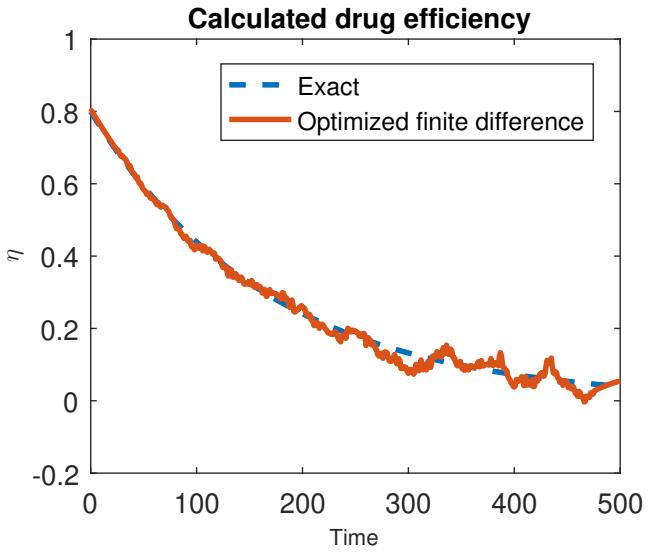
$$\begin{cases} \dot{u}_1 = s - k u_1 u_4 - \mu u_1 + (\eta\alpha + b)u_2, \\ \dot{u}_2 = k u_1 u_4 - (\mu_1 + \alpha + b)u_2, \\ \dot{u}_3 = (1 - \eta)\alpha u_2 - \delta u_3, \longleftarrow \\ \dot{u}_4 = N\delta u_3 - c u_4. \end{cases} \tag{4}$$

From the third equation we can get:

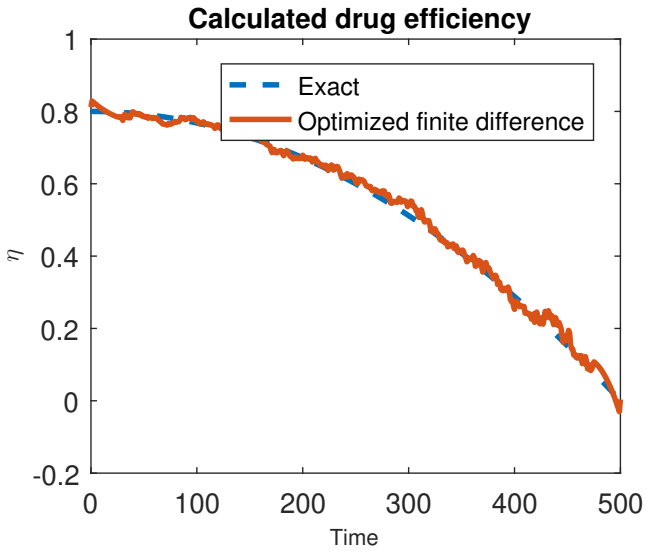
$$\eta(t) = 1 - \frac{\dot{u}_3 + \delta u_3}{\alpha u_2}. \tag{5}$$

The derivative $\dot{u}_3$ can be approximated using the central difference rule $\dot{u}_3(t) \approx \frac{u_3(t+\tau) - u_3(t-\tau)}{2\tau}$, τ is the time step.

Continuous observation on entire time domain



$\eta(t) = 0.8e^{-3t/T}, \sigma = 15\%$



$\eta(t) = 0.8(1 - t/T), \sigma = 20\%$

The random noise:

$$u_\sigma(t) = u_\sigma(t)(1 + \sigma\alpha),$$

where $\sigma \in [0, 1]$ is nose level and $\alpha \in [-1, 1]$ is random number.

[ME] M. Eriksson, Parameter identification in a mathematical model of HIV infection with drug therapy, Master thesis, https://gupea.ub.gu.se/handle/2077/54664, 2017.

Theory: Ill-posed and inverse problems

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces. Let $G \subseteq \mathcal{H}_1$. Consider a continuous mapping $F : G \to \mathcal{H}_2$ and let $y \in \mathcal{H}_2$ be given. Suppose we want to find $x \in G$ such that

$$F(x) = y. \tag{6}$$

Definition: The inverse problem is well-posed by Hadamard if

1. For each $y \in \mathcal{H}_2$ there exists an $x \in G$ that solves the problem above (i.e. F is onto).
2. For each $y \in \mathcal{H}_2$ there is at most one such x (i.e. F is one-to-one).
3. The solutions $x(y)$ depends continuously on y (i.e. the inverse mapping F^{-1} is continuous).

M. M. Lavrentiev *Some Improperly Posed Problems of Mathematical Physics*, Springer Tracts in Natural Philosophy, vol. 11, Springer Verlag, Berlin, 1967.

Theory: Stable solution of ill-posed problems

- **Assume:** Forward problem is well-posed (but the inverse problem is ill-posed). There exists an ideal exact solution, x^*, with a small neighborhood, G, where the inverse problem is well-posed (!). The set G is a compact set.
- Then given noisy observations, y_δ, the best approximation is obtained by finding $x \in G$ that minimizes

$$Q(x) = \|F(x) - y_\delta\|^2. \tag{7}$$

This is called the *quasi-solution* or *least squares solution*.

But

- In general, we do not know x^* and thus usually not the set G.
- The minimum may not be unique. There may be local minima and ravines that make minimization difficult.

Theory: Tikhonov functional for model problem

The Tikhonov functional $J_\gamma : G \to \mathbb{R}$ is defined as

$$J_\gamma(x) = \frac{1}{2}\|F(x) - y_\delta\|^2 + \frac{1}{2}\gamma\|x - x^0\|^2, \tag{8}$$

where $x^0 \in G$, $\gamma(\delta) \to 0$ and $\frac{\delta^2}{\gamma(\delta)} \to 0$ as $\delta \to 0$.

- The Tikhonov functional is *Fréchet differentiable* and *locally strongly convex* in a neighborhood of its minimum if $\|x^0 - x^*\|$ is small enough.
- If $\|x^0 - x^*\|$ is small enough, then the regularized solution x_γ of (7) will be also in this neighborhood such that $\|x_\gamma - x^*\| \leq \xi\|x^0 - x^*\|, \xi \in (0, 1)$, i.e. the minimum of (8) is always a better approximation to x^* than x^0.

[BK] A. B. Bakushinskii and M. Yu. Kokurin, *Iterative Methods for Approximate Solution of Inverse Problems*, Springer, New York, 2004.

The parameter identification problem

Parameter Identification Problem (PIP). Assume that parameters $\{s, \mu, k, \mu_1, \alpha, b, \delta, c, N\}$ are known. Assume further that the function $\eta(t)$ is unknown inside the domain Ω_T. The PIP is: determine $\eta(t)$ for $t \in \Omega_T$, under the condition that the virus population function $g(t)$ is known

$$u_4(t) = g(t), \quad t \in [T_1, T_2], 0 \leq T_1 < T_2 \leq T. \tag{9}$$

Here, the function $g(t)$ presents observations of the function $u_4(t)$ inside the observation interval $[T_1, T_2]$.
The Tikhonov functional is given by

$$J(\eta) = \frac{1}{2}\int_{T_1}^{T_2}(u_4(t) - g(t))^2 z_\zeta(t)dt + \frac{1}{2}\gamma\int_0^T(\eta - \eta^0)^2 dt, \tag{10}$$

where g_i are observations, η^0 the initial guess, γ the regularization parameter and z_ζ a bump function making J continuous.

[BG] L. Beilina, I. Gainova, *Time-adaptive FEM for distributed parameter identification in mathematical model of HIV infection with drug therapy*, Inverse Problems and Application, Springer Proceedings in Mathematics and Statistics, vol. 120, Springer, 2015, pp. 111-124.

Optimization method

- We seek for a stationary point of the Tikhonov functional with respect to η which satisfies

$$J'(\eta)(\bar{\eta}) = 0, \quad \forall \bar{\eta} \in H. \tag{11}$$

- To do this, we introduce the Lagrangian

$$L(v) = J(\eta) + \sum_{i=1}^{4} \int_0^T \lambda_i(\dot{u}_i - f_i)dt. \tag{12}$$

where the f_i are the RHS of (1) and $v = (u, \lambda, \eta) \in U$, where

$$\begin{aligned} H_u^1(\Omega_T) &= \{f \in H^1(\Omega_T) : f(0) = 0\}, \\ H_\lambda^1(\Omega_T) &= \{f \in H^1(\Omega_T) : f(T) = 0\}. \\ U &= H_u^1(\Omega_T) \times H_\lambda^1(\Omega_T) \times C(\Omega_T). \end{aligned} \tag{13}$$

- Now we seek

$$0 = L'(v)(\bar{v}) = \frac{\partial L}{\partial \lambda}(v)(\bar{\lambda}) + \frac{\partial L}{\partial u}(v)(\bar{u}) + \frac{\partial L}{\partial \eta}(v)(\bar{\eta}), \quad \forall \bar{v} = (\bar{\lambda}, \bar{u}, \bar{\eta}) \in U. \tag{14}$$

Finite element formulation

We discretize the initial Ω_T into a uniform mesh, J_τ, with step length $\tau := t_k - t_{k-1}$ and time intervals $J_k = (t_{k-1}, t_k]$. Then, we define the following spaces:

$$\begin{aligned} W_\tau^u(\Omega_T) &= \{f \in H_u^1 : f|_{J_k} \in P^1(J_k)\ \forall J_k \in J_\tau\}, \\ W_\tau^\lambda(\Omega_T) &= \{f \in H_\lambda^1 : f|_{J_k} \in P^1(J_k)\ \forall J_k \in J_\tau\}, \\ W_\tau^\eta(\Omega_T) &= \{f \in L_2(\Omega_T) : f|_{J_k} \in P^0(J_k)\ \forall J_k \in J_\tau\}, \\ U_\tau &= W_\tau^u(\Omega_T) \times W_\tau^\lambda(\Omega_T) \times W_\tau^\eta(\Omega_T). \end{aligned} \tag{15}$$

The finite element method for (14) is: find $v_\tau \in U_\tau$ such that

$$L'(v_\tau; \bar{v}) = 0, \quad \forall \bar{v} \in U_\tau. \tag{16}$$

Newton's method is used for the solution of forward and adjoint problems.

Newton's method for the forward problem

For the discretization

$$\frac{\partial u}{\partial t} = \frac{u^{k+1} - u^k}{\tau_k}$$

the variational formulation of the forward problem for all $\bar{u} \in H_u^1(\Omega_T)$ is:

$$(u^{k+1}, \bar{u}) - (u^k, \bar{u}) - \tau_k (f(u^{k+1}), \bar{u}) = 0. \tag{17}$$

Denoting

$$\begin{aligned} \tilde{u} &= u^{k+1}, \\ F(\tilde{u}) &= \tilde{u} - \tau_k f(\tilde{u}) - u^k \end{aligned} \tag{18}$$

we can rewrite (17) as

$$(F(\tilde{u}), \bar{u}) = 0. \tag{19}$$

For solution $F(\tilde{u}) = 0$ the Newton's method can be used:

$$\tilde{u}^{n+1} = \tilde{u}^n - [F'(\tilde{u}^n)]^{-1} \cdot F(\tilde{u}^n). \tag{20}$$

Here, we can determine $F'(\tilde{u}^n)$ via definition of $F(\tilde{u})$ in (18) as

$$F'(\tilde{u}^n) = I - \tau_k f'(\tilde{u}^n).$$

Adaptive time-mesh refinement

Refinement of the time mesh is be based on the following theorem [BG].

Theorem (A posteriori error estimate for the regularized solution)

Let $\eta_\tau \in W_\tau^\eta$ be a finite element approximation on the finite element mesh J_τ of the minimizer $\eta \in L^2(\Omega_T)$ with the mesh function $\tau(t)$. Then there exists an interpolation constant C_I independent on τ such that the following a posteriori error estimate for the minimizer η holds

$$\|\eta_\tau - \eta\|_{L_2(\Omega_T)} \leq \frac{\|R(\eta_\tau)\|}{\gamma} C_I \|\tau \eta_\tau\|_{L_2(\Omega_T)} \ \forall \eta_\tau \in W_\tau^\eta, \tag{21}$$

where $R(\eta_\tau)$ is the residual defined as

$$R(\eta_\tau)(t) = \gamma(\eta_\tau - \eta^0)(t) + \alpha u_{2\tau}(\lambda_{3\tau} - \lambda_{1\tau})(t). \tag{22}$$

[BG] L. Beilina, I. Gainova, *Time-adaptive optimization in a parameter identification problem of HIV infection*, arXiv:1912.01112v1

Conjugate gradient algorithm

Algorithm

- Step 0. Choose time partition J_τ of the time interval $(0, T)$. Start with the initial guess $\eta = \eta^0$ and compute the sequences of η^m via the following steps:
- Step 1. Compute solutions $u(t, \eta)$ and $\lambda(t, \eta)$ of state and adjoint problems on J_τ.
- Step 2. Update the coefficient $\eta := \eta^{m+1}$ on J_τ using the conjugate gradient method

$$\begin{aligned} \eta^{m+1} &= \eta^m + \alpha^m d^m(x), \\ d^m(x) &= -g^m(x) + \beta^m d^{m-1}(x), \\ \beta^m &= \frac{\|g^m(x)\|^2}{\|g^{m-1}(x)\|^2}, \end{aligned} \tag{23}$$

where $d^0(x) = -g^0(x)$. In (23) the step size α in the gradient update is computed as

$$\alpha^m = -\frac{(g^m, d^m)}{\gamma^m \|d^m\|^2}. \tag{24}$$

Conjugate gradient algorithm

Algorithm

The regularization parameter γ is computed iteratively accordingly to [BKS] as

$$\gamma^m = \frac{\gamma_0}{(m+1)^p}, p \in (0,1). \tag{25}$$

- Step 3. Stop computing η^m and obtain the function η at $M = m$ if either $\|g^m\|_{L_2(U_\tau)} \leq \theta$ or norms $\|\eta^m\|_{L_2(U_\tau)}$ are stabilized. Here θ is the tolerance in updates m of gradient method. Otherwise set $m := m + 1$ and go to step 1.
- Refine the time mesh where

$$|R^M(\eta_\tau)(t)| \geq \beta \max_{t \in \Omega_T} |R^M(\eta_\tau)(t)|, \tag{26}$$

where $\beta \in (0,1)$ is chosen by the user. Then go to step 1.

[BKS] A. Bakushinsky, M.Y. Kokurin, A. Smirnova, *Iterative Methods for Ill-posed Problems*, Inverse and Ill-Posed Problems Series 54, De Gruyter, 2011.

Time-adaptive two-stage procedure for random data

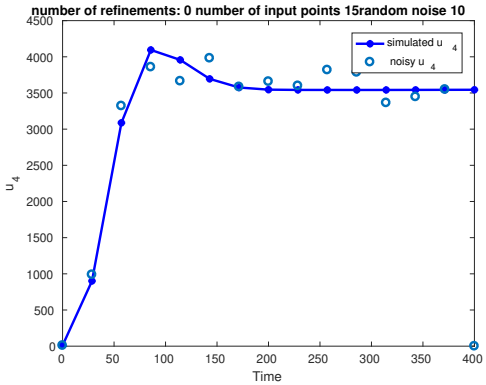
- 0. Initialize data g on the initial time-mesh J_τ.
- 1. First stage: obtain initial guess η_0 by solving least squares problem $\min_\eta \|A\eta - g\|_2^2$. Can be used method of normal equations, QR or SVD decompositions. As test functions in constructing of elements of matrix A can be used $t^d, d = 1, ..., m, t \in [0, T]$, or splines.
- 2. Second stage: minimize the Tikhonov functional

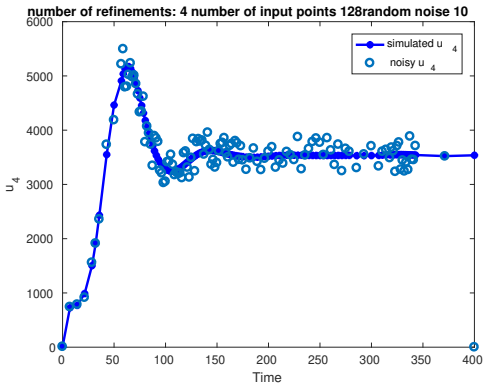
$$J(\eta) = \frac{1}{2}\int_{T_1}^{T_2}(u_4(t) - g(t))^2 z_\zeta(t)dt + \frac{1}{2}\gamma\int_0^T(\eta - \eta^0)^2 dt, \qquad (27)$$

  where g_i are observations, η^0 the initial guess obtained at the first stage, γ the regularization parameter and z_ζ a bump function making J continuous.
- 3. Refine mesh using a posteriori error indicator and obtain a new time-mesh J_τ. Interpolate observed data g into a new time mesh. Go to step 1.

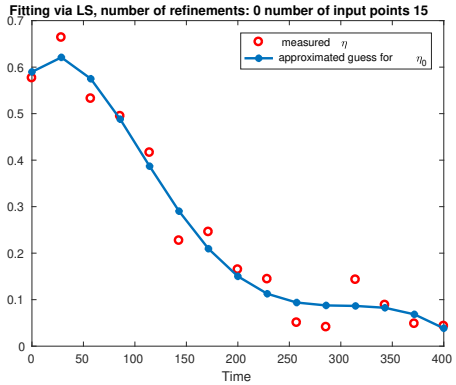
Tests with observed virus function u_4 on $T_{obs} = [0, 400]$

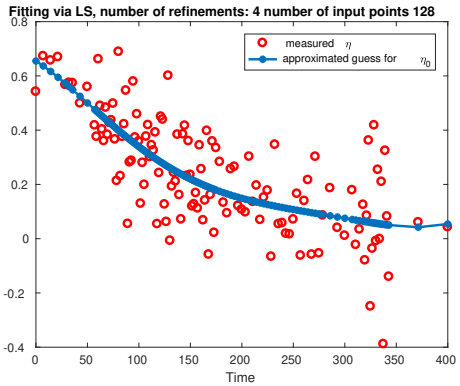
$u_4, \sigma = 10\%$



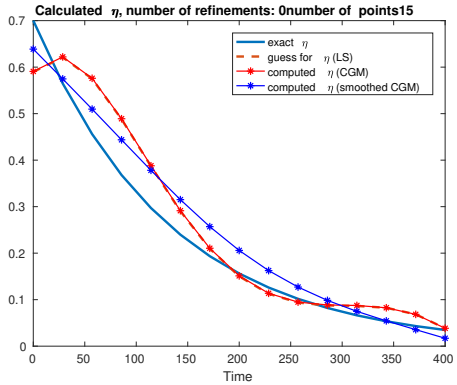


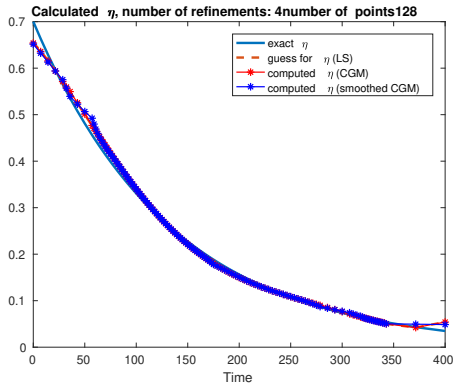
LS fitting to η_0





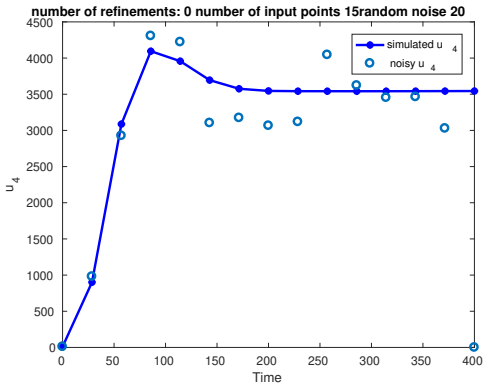
optimized η



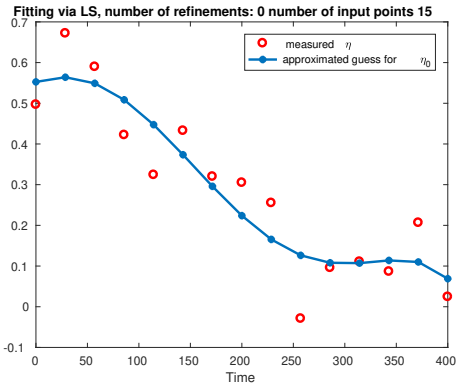


Tests with observed virus function u_4 on $T_{obs} = [0, 400]$

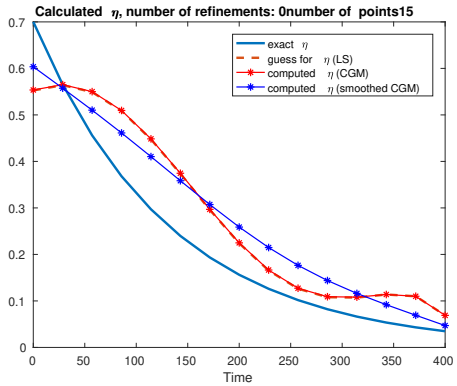
$u_4, \sigma = 20\%$

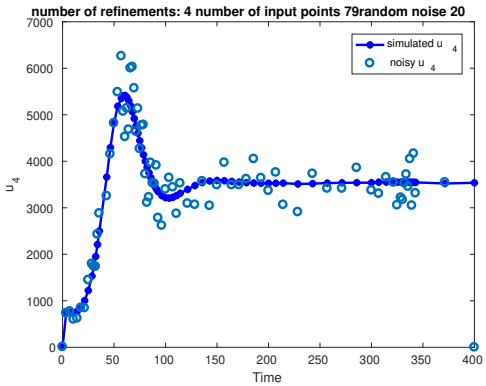


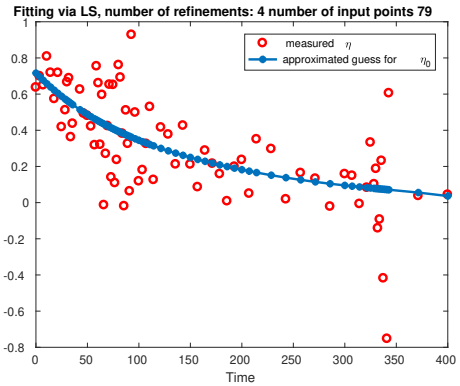
LS fitting to η_0

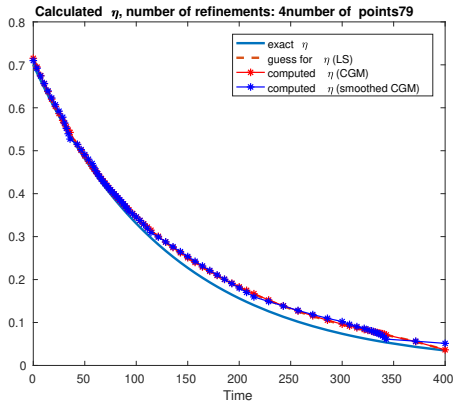


optimized η



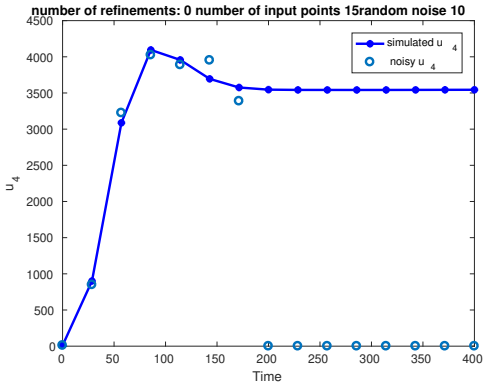


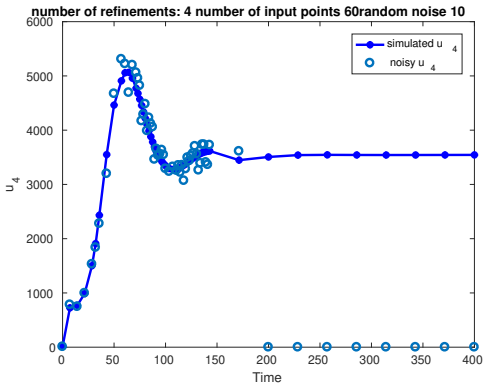




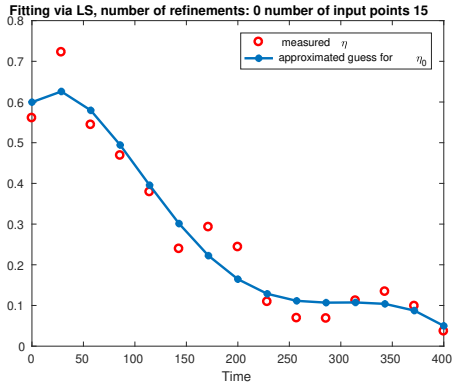
Tests with observed virus function u_4 on $T_{obs} = [0, 200]$

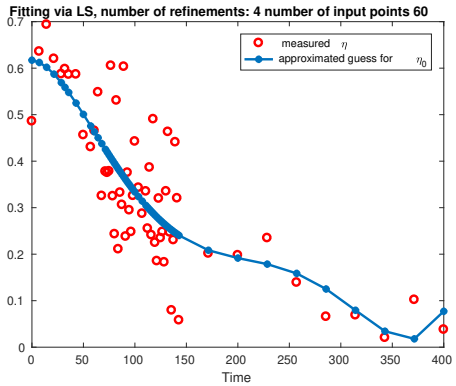
$u_4, \sigma = 10\%$



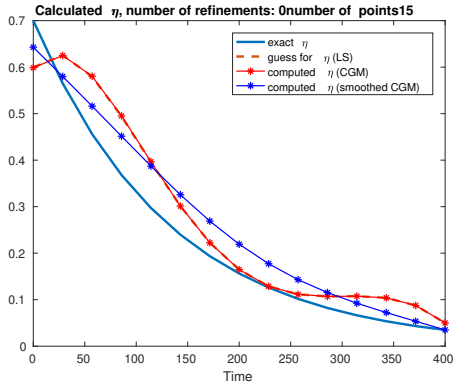


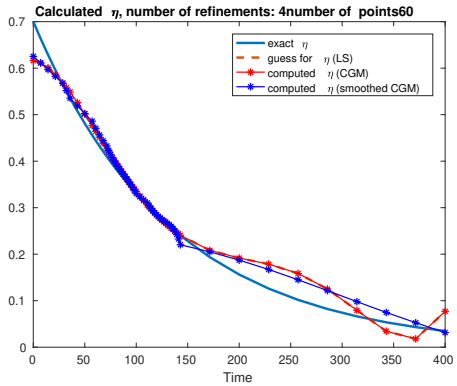
LS fitting to η_0





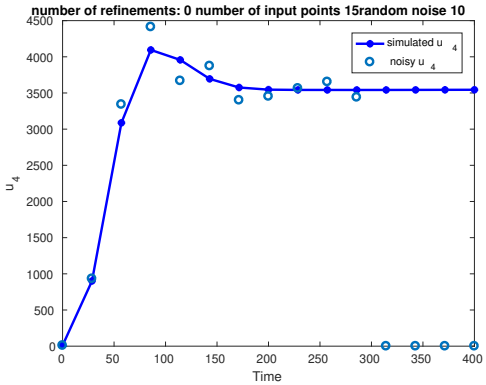
optimized η

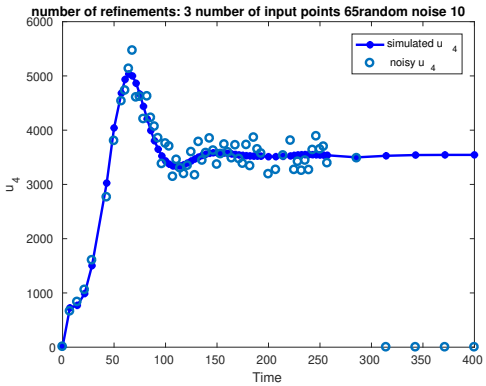




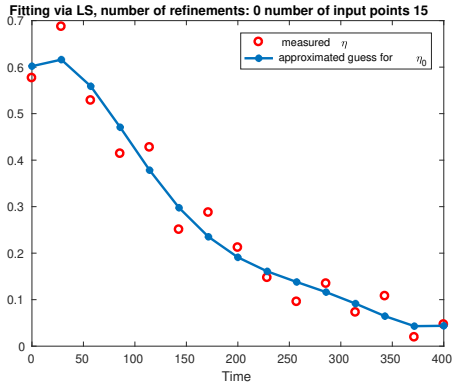
Tests with observed virus function u_4 on $T_{obs} = [0, 300]$

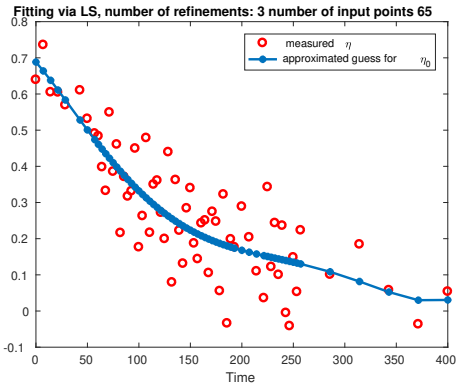
$u_4, \sigma = 10\%$



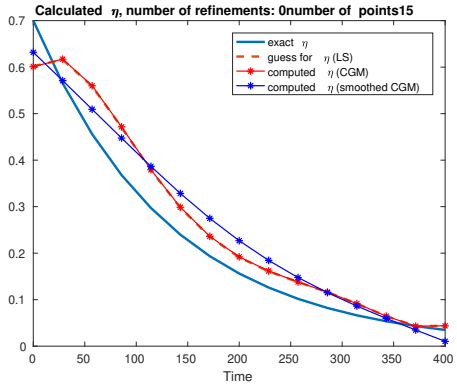


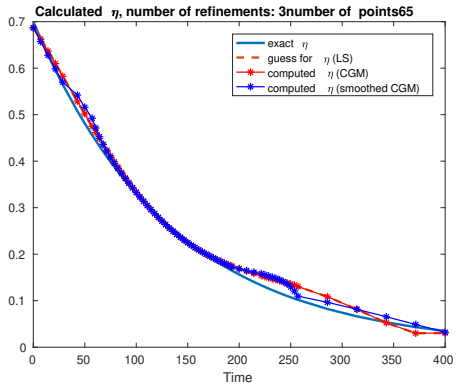
LS fitting to η_0





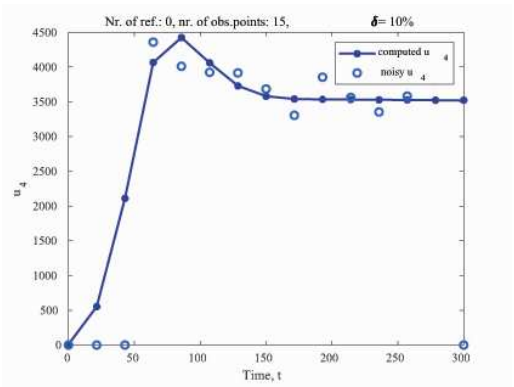
optimized η



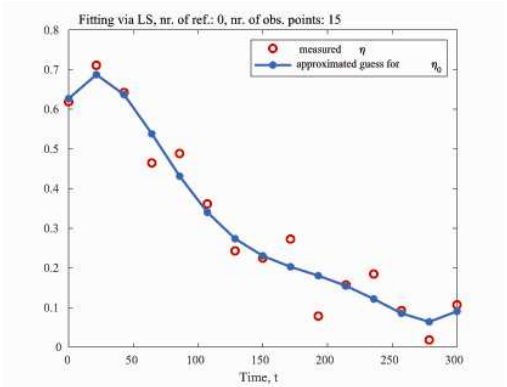


Tests with observed virus function u_4 on $T_{obs} = [50, 300]$

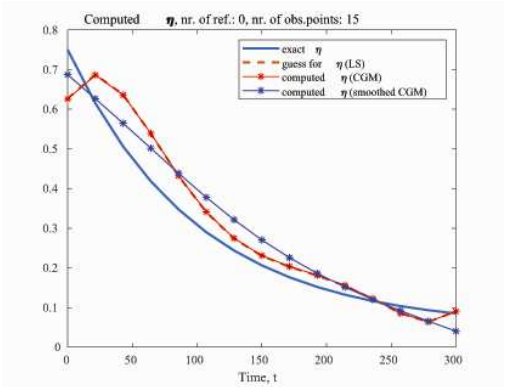
$u_4, \sigma = 10\%$

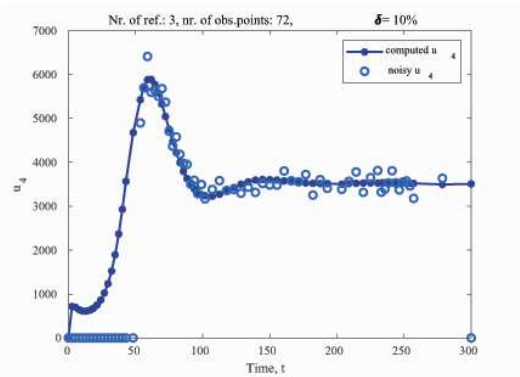


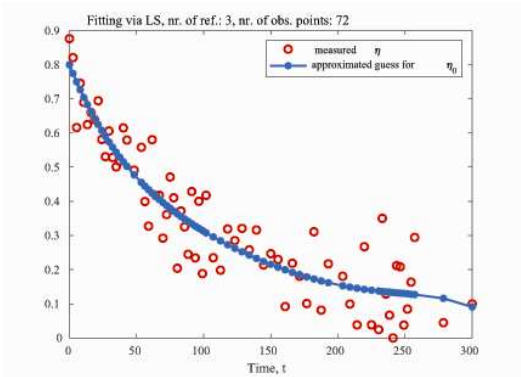
LS fitting to η_0

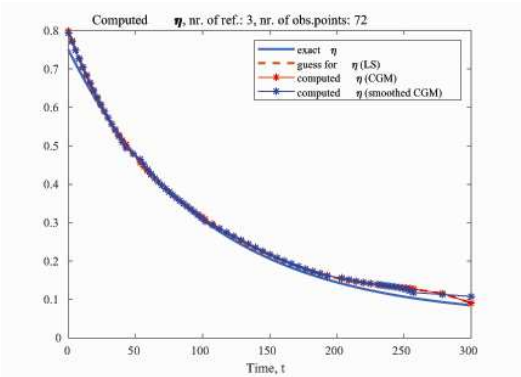


optimized η



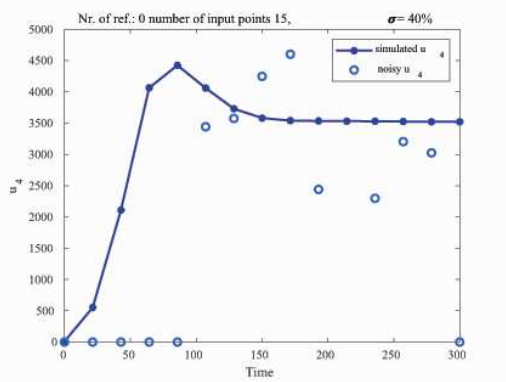


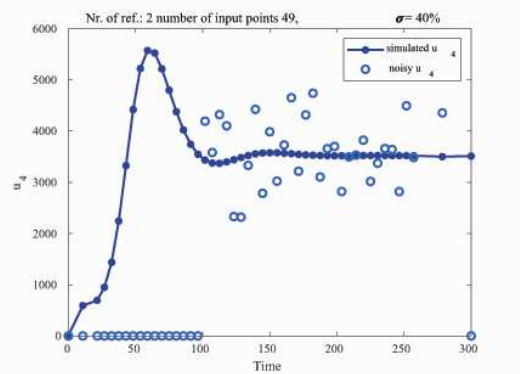




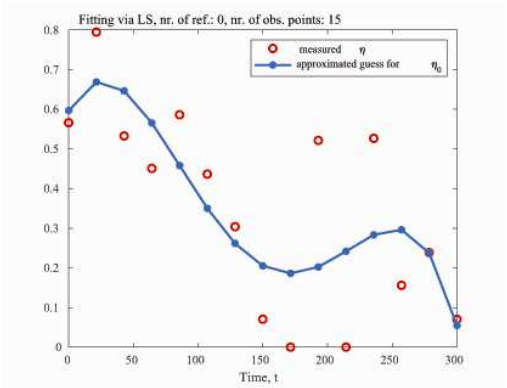
Tests with observed virus function u_4 on $T_{obs} = [100, 300]$

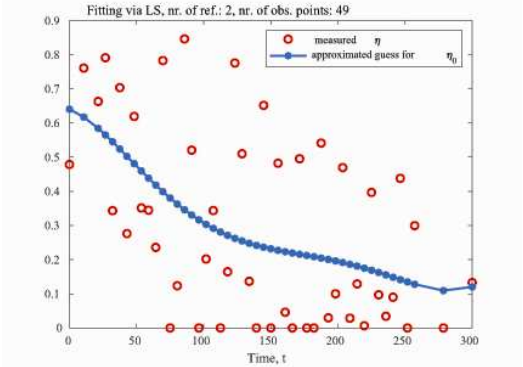
$u_4, \sigma = 40\%$



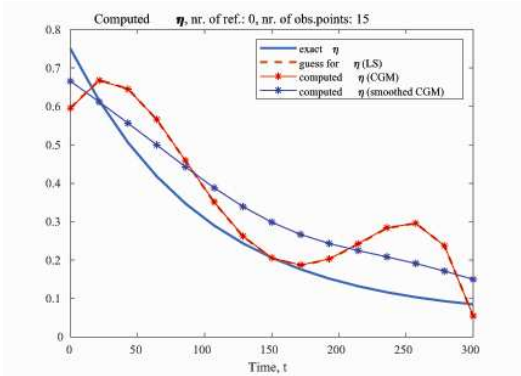


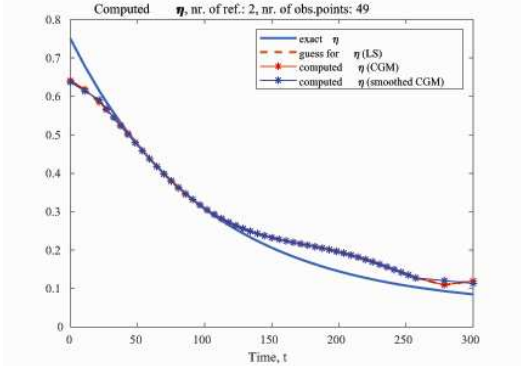
LS fitting to η_0





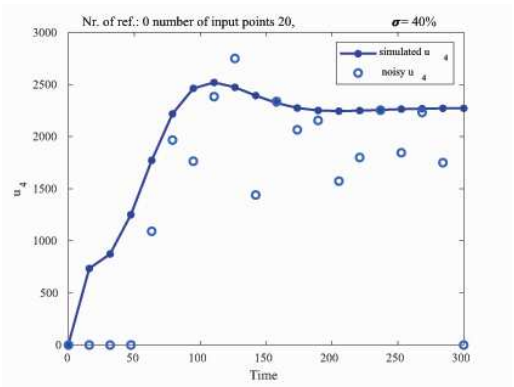
optimized η

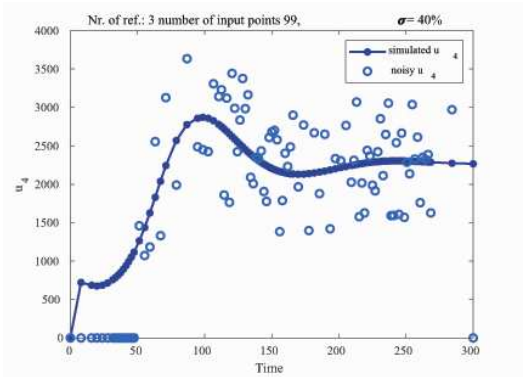




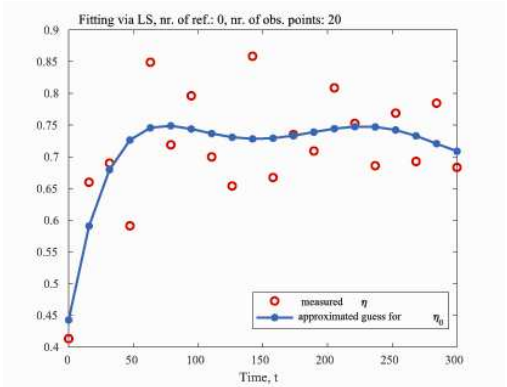
Tests with observed virus function u_4 on $T_{obs} = [50, 300]$

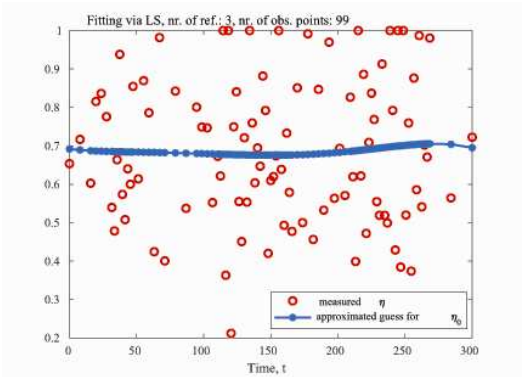
$u_4, \sigma = 40\%$



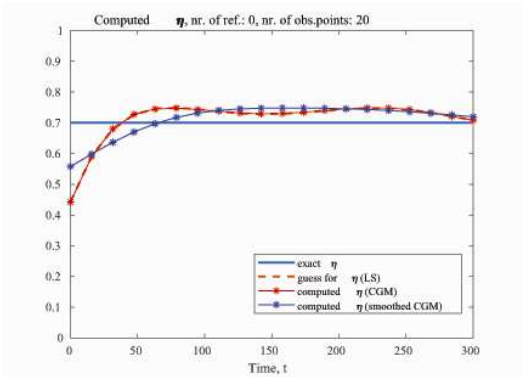


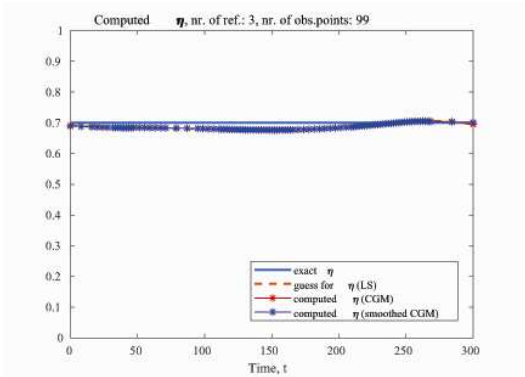
LS fitting to η_0





optimized η





Conclusions

- The adaptive optimization method works for simulated data. It is desirable to test method on real data. How to get them ?
- Test other functionals can be minimized for the same model. We minimized the functional

$$J(\eta) = \frac{1}{2}\int_{T_1}^{T_2}(u_4(t) - g_i(t))^2 z_\zeta(t)dt + \frac{1}{2}\gamma\int_0^T(\eta - \eta^0)^2 dt. \quad (28)$$

- One can minimize other functionals, see for example,

  E. F. Arruda, C. M. Dias, , C. V. de Magalhaes, D. H. Pastore, R. C. A. Thomé, H. M. Yang, An optimal control approach to HIV immunology, *Applied Mathematics*, 1115-1130, 2015. http://dx.doi.org/10.4236/am.2015.66102

- The method can be extended to other biological models and to the reconstruction of several unknown parameters.