



## **Time-Adaptive Determination of Drug Efficacy in Mathematical Model of HIV Infection**

Downloaded from: <https://research.chalmers.se>, 2024-04-24 22:22 UTC

Citation for the original published paper (version of record):

Beilina, L., Eriksson, M., Gainova, I. (2024). Time-Adaptive Determination of Drug Efficacy in Mathematical Model of HIV Infection. *Differential Equations and Dynamical Systems*, 32(1): 313-347. <http://dx.doi.org/10.1007/s12591-021-00572-w>

N.B. When citing this work, cite the original published paper.



# Time-Adaptive Determination of Drug Efficacy in Mathematical Model of HIV Infection

L. Beilina<sup>1</sup> · M. Eriksson<sup>2</sup> · I. Gainova<sup>3</sup>

Accepted: 25 May 2021  
© The Author(s) 2021

## Abstract

The paper considers a time-adaptive finite element method for determination of drug efficacy in a parameter identification problem (PIP) for a system of ordinary differential equations (ODE) that describes dynamics of the primary human immunodeficiency virus (HIV) infection with drug therapy. Tikhonov's regularization method, optimization approach and finite element method to solve this problem are presented. A posteriori error estimates in the Tikhonov's functional and reconstructed parameter are derived. Based on these estimates a time adaptive algorithm is formulated and numerically tested for different scenarios of noisy observations of virus population function. Numerical results show a significant improvement of reconstruction of drug efficacy parameter using the local time-adaptive mesh refinement method compared to the gradient method applied on a uniform time mesh.

## Introduction

Mathematical modeling of the immune processes is an essential part of the research in immunology [15, 16, 22]. Despite the emergence of a great amount of new high-performance methods for experimental analysis of the immunity, the results of mathematical modeling are relevant, in particular, in clinical practice in order to work out optimal, individually customized strategies for treatment of the pathological process (bacterial/viral infections or tumor growth).

It is well known that physiological parameters vary between individual patients and thus, personalized treatment approaches require the development of robust and efficient

---

✉ L. Beilina  
larisa@chalmers.se

M. Eriksson  
martin.eriksson.2@gu.se

I. Gainova  
gajnova@math.nsc.ru

<sup>1</sup> Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 42196 Göteborg, Sweden

<sup>2</sup> Department of Marine Sciences, University of Gothenburg, 40530 Göteborg, Sweden

<sup>3</sup> Sobolev Institute of Mathematics, Novosibirsk 630090, Russia

parameter estimation methods to assimilate individual data with mathematical models [6]. The problems of parameter identification in mathematical models are often non-linear and ill-posed, and thus challenging to solve numerically [45, 46]. The computational algorithms for parameter identification that we work on, are based on an adaptive time-mesh refinement [13] for coefficient inverse problems (CIPs). The main idea of our work is adoption of results for space mesh refinement developed in [13] for solution of hyperbolic CIPs, to the parameter identification problem (PIP) for reconstruction of parameters on the time mesh. More precisely, first, we determine candidate parameter at known initial (coarse) time partition. Then we refine time-mesh locally only at a such time intervals where a posteriori error indicator is large and compute new time-dependent control functions on a new time mesh until the error in the computed residual is reduced to the desired accuracy. The adaptive finite element method has shown that it significantly improves reconstruction of parameters when solving the coefficient inverse problems for hyperbolic PDE [5, 7, 8, 12, 13].

We note that the main goal of our work is to present mathematical framework of a posteriori error estimation for solution of PIP's and to show usefulness of the time-adaptive error control for determination of parameters in PIP which we demonstrate on the example of the model of HIV infection. More than 35 years have passed since the discovery of the etiological agent of AIDS—human immunodeficiency virus (HIV), nevertheless, the problem of the spread of HIV infection, treatment and quality of life for people living with HIV still remains relevant: the number of newly infected does not decrease. Appearing of highly active antiretroviral therapy (HAART) in 1996-1997 has led to a significant improvement in the quality of life of patients, has caused a clear decrease in AIDS-related diseases and mortality. HAART provides treatment protocols with using combinations (two or more) of antiviral drugs which affect both the different stages of viral replication and prevent HIV from entering the host cell.

We took the simplified model of HIV infection proposed in [42] as an illustrative example to show effectiveness and robustness of identification time-dependent parameter drug efficacy using a local time-mesh refinement algorithm. This work is a continuation of the works by authors [9, 10] where was introduced the time-adaptive finite element method for parameter identification problems. Compared to other optimal control algorithms for solution of PIP, see, for example, works [1, 2, 26] and references therein, our time-adaptive algorithms are based on rigorous finite element a posteriori error analysis for the error in the Tikhonov's functional or for the error in the reconstructed parameter. The same approach can be applied to the solution of any other PIP and particularly, for more complicated models of HIV infection which involve more unknown functions and parameters [3, 27, 34–36, 40, 41, 48]. However, these models are much complicated compared with studied here model of [42], and can be considered as topic for a future research.

As was mentioned above, in [10] was studied the optimal control problem of reconstruction of drug efficacy in the model of HIV infection when measurements of all functions in time of the model ODE system were used what, actually, is not realistic problem. Moreover, numerical simulations were not presented in [10]. In the current work we fill this gap and study a more realistic case when instead of measuring all four functions in ODE system, we take measurements only of the virus population function. New a posteriori error estimate between regularized and computed drug efficacy is derived. Based on this estimate, a time adaptive algorithm is formulated and numerically tested on the optimal determination of drug efficacy in time domain from noisy measurements of virus population function. Extended numerical studies for different noise levels in the virus population function and for different values of time for initial observation of this function are presented in

the preprint version of this work [11]. We note that compared to the current work proofs of Theorems 1, 2, and 3 are not presented in [11], as well as study on stability of solution of the forward problem and study of ill-posedness of the PIP problem.

The time-adaptive method proposed in this paper can eventually be used by clinicians to determine the drug-response for each treated individual. The exact knowledge of the personal drug efficacy can aid in the determination of the most suitable drug as well as the most optimal dose to an individual, in the long run resulting in a personalized treatment with maximum efficacy and minimum adverse drug reactions.

The outline of the paper is as follows. The biological description of the mathematical model is given in “[The Mathematical Model and Its Biological Description](#)”. “[Inverse Problems and Ill-Posedness](#)” is based on material of the Master’s thesis [25] and discusses ill-posedness of parameter identification problems and Tikhonov’s regularization method for their solutions. In “[The Parameter Identification Problem](#)” the parameter identification problem is formulated. The optimization method to solve the parameter identification problem is presented in “[Optimization Method](#)”. The finite element method is formulated in “[Finite Element Discretization](#)” and a posteriori error estimates are derived in “[A Posteriori Error Estimates](#)”. An adaptive algorithm for solution of PIP is presented in “[Algorithms for Solution of PIP](#)”. Finally, in “[Numerical Results](#)” numerical examples illustrate effectiveness of the proposed time-adaptive algorithm.

## The Mathematical Model and Its Biological Description

The main cellular targets of HIV are the immune system cells that have CD4 receptors at their surface, called *CD4+ T-cells*. HIV differs from other viruses by a high mutation level, the mutation rate is  $10^{-5}$ – $10^{-4}$  per nucleotide during one replication cycle [30]. High genetic variability allows HIV to skillfully avoid humoral and cellular defense factors and the effects of drugs. The constant presence of large reservoirs of latently infected cells is one of the important feature of the pathogenesis of HIV infection. Another paradoxical feature of HIV is that activation of the immune system does not lead to a suppression of virus multiplication, but rather to activation of latently infected cells, which start to produce new viral particles. These factors are the major obstacles for antiviral therapy and the development of efficient vaccines [18, 19]. According to the latest data on HIV (UNAIDS, 2020) [47], there are currently 38 million people living with HIV globally and 1.7 million people became newly infected with HIV in 2019.

The HIV life cycle starts with attachment of the viral envelope protein gp120 to the cell surface via interaction with CD4 receptor. In general case for living organisms the genetic information goes from the storage in DNA through messenger RNA (mRNA) to protein synthesis in the ribosomes. The process of converting the genetic information from DNA to mRNA is called *transcription* [25]. In the case of retroviruses, such as HIV, HIV’s genetic information is encoded in form of RNA. Having fused with the cell membrane, HIV releases its genetic material (viral RNA) and enzymes into the CD4+ T-cell. Here viral RNA is reversely transcribed into HIV DNA, which is compatible with genetic material of the host cell [*reverse transcription* (RT)]. To perform the reverse transcription of RNA into DNA, HIV carries its own enzyme called *reverse transcriptase*. Viral DNA is transported to the cell’s nucleus and incorporated into the DNA of the host cell (*integration*). This process is made possible by the enzyme *integrase*. The individual components of HIV are then produced within the CD4+ T-cell.

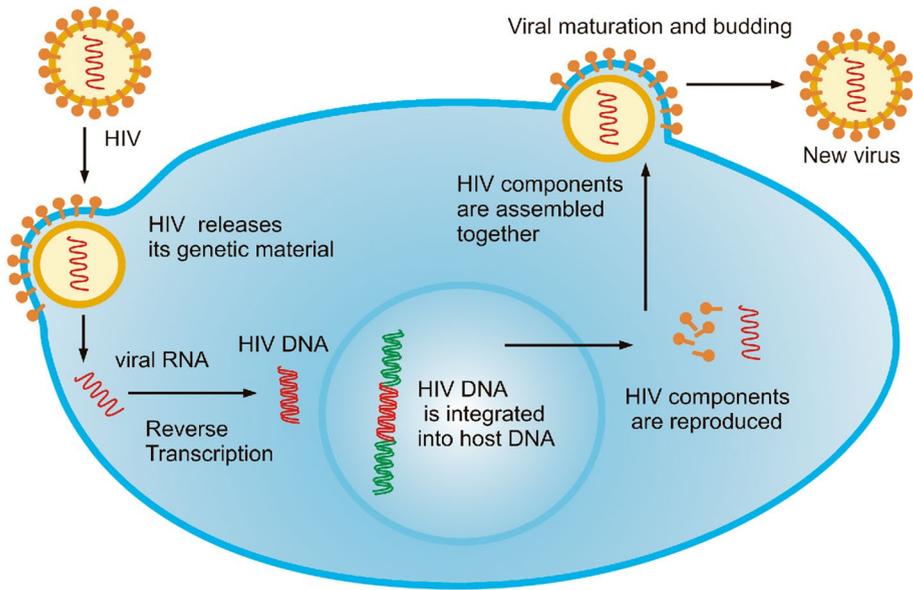


Fig. 1 The HIV life cycle

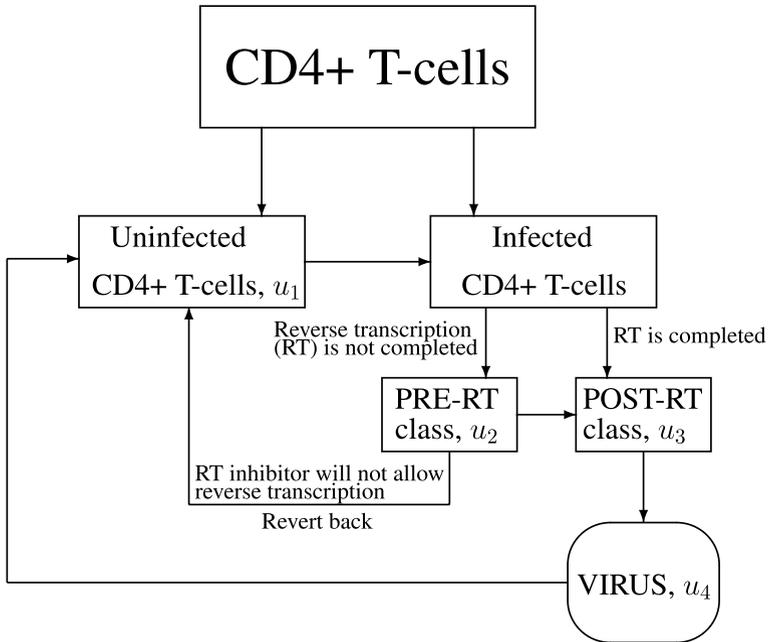
The individual components of HIV are then assembled together to make new HIV viruses. This process depends on the enzyme *protease*. The newly matured HIV particles are released from the CD4+ T-cell. These are ready to infect other CD4+ T-cells and begin the replication process all over again. The process of HIV life cycle described above is illustrated in Fig. 1. Antiretroviral drugs blocking the enzyme reverse transcriptase (called *Reverse Transcriptase Inhibitors*) will be able to prevent the production of new viruses [18, 37].

Our basic mathematical model in this work is the model proposed in [42] which describes the effect of Reverse Transcriptase Inhibitor (RTI) on the dynamics of HIV infection. In this model the infected class of CD4+ T-cells is subdivided into two subclasses: pre-RT class and post-RT class. Pre-RT class consists of the infected CD4+ T-cells in which reverse transcription is not completed, and post-RT class consists of those infected CD4+ T-cells where the reverse transcription is completed such that they are capable to produce virus.

Throughout the paper we denote by  $\Omega_T = [0, T]$  the time domain for  $T > 0$ , where  $T$  is the final observation time. The mathematical model which we use in this note, is:

$$\begin{cases} \frac{du_1}{dt} = f_1(u(t), \eta(t)) = s - ku_1(t)u_4(t) - \mu u_1(t) + (\eta(t)\alpha + b)u_2(t), \\ \frac{du_2}{dt} = f_2(u(t), \eta(t)) = ku_1(t)u_4(t) - (\mu_1 + \alpha + b)u_2(t), \\ \frac{du_3}{dt} = f_3(u(t), \eta(t)) = (1 - \eta(t))\alpha u_2(t) - \delta u_3(t), \\ \frac{du_4}{dt} = f_4(u(t), \eta(t)) = N\delta u_3(t) - cu_4(t), \end{cases} \quad (1)$$

with initial conditions



**Fig. 2** The effect of reverse transcriptase inhibitor (RTI) on the dynamics of HIV infection described by the model problem (1)

$$\begin{aligned}
 u_1(0) = u_1^0 = 300 \text{ mm}^{-3}, \quad u_2(0) = u_2^0 = 10 \text{ mm}^{-3}, \\
 u_3(0) = u_3^0 = 10 \text{ mm}^{-3}, \quad u_4(0) = u_4^0 = 10 \text{ mm}^{-3}.
 \end{aligned}
 \tag{2}$$

In system (1) the functions are defined as follows:

- $u_1(t)$  – uninfected target cells population
- $u_2(t)$  – infected target cells before Reverse Transcription (pre-RT class)
- $u_3(t)$  – infected target cells in which Reverse Transcription is completed and they are capable of producing virus (post-RT class)
- $u_4(t)$  is the virus population function.

The parameters which are used in system (1) are described in Table 1 and are taken from the specialized literature [32, 33]. The initial data (2) is chosen such that they satisfy two steady states which we discuss in “Existence and Lyapunov Stability of Steady State”. Figure 2 illustrates the effect of Reverse Transcriptase Inhibitor (RTI) on the dynamics of HIV infection for the mathematical model (1).

The system (1) can be presented in the following compact form:

$$\begin{cases} \frac{du}{dt} = f(u(t), \eta(t)) & t \in [0, T], \\ u(0) = u^0, \end{cases}
 \tag{3}$$

where we have denoted all involved functions as

**Table 1** Parameters dataset

Parameter	Value	Units	Description
$s$	10	$\text{mm}^{-3} \text{day}^{-1}$	Inflow rate of T cells
$\mu$	0.01	$\text{day}^{-1}$	Natural death rate of T cells
$k$	$2.4\text{E}-5$	$\text{mm}^3\text{day}^{-1}$	Interaction-infection rate of T cells
$\mu_1$	0.015	$\text{day}^{-1}$	Death rate of infected cells
$\alpha$	0.4	$\text{day}^{-1}$	Transition rate from pre-RT infected T cells class to post-RT class
$b$	0.05	$\text{day}^{-1}$	Reverting rate of infected cells return to uninfected class
$\delta$	0.26	$\text{day}^{-1}$	Death rate of actively infected cells
$c$	2.4	$\text{day}^{-1}$	Clearance rate of virus
$N$	1000	vir/cell	Total number of viral particles produced by an infected cell
$\eta$	[0, 1]	NA	Control, dosage of the reverse transcriptase inhibitor (drug efficacy)

$$\begin{aligned}
 u &= u(t) = (u_1(t), u_2(t), u_3(t), u_4(t))^T, \\
 u^0 &= (u_1(0), u_2(0), u_3(0), u_4(0))^T, \\
 \frac{du}{dt} &= \left( \frac{\partial u_1}{\partial t}, \frac{\partial u_2}{\partial t}, \frac{\partial u_3}{\partial t}, \frac{\partial u_4}{\partial t} \right)^T, \\
 f(u(t), \eta(t)) &= (f_1, f_2, f_3, f_4)^T(u(t), \eta(t)) \\
 &= (f_1(u_1, \dots, u_4, \eta(t)), \dots, f_4(u_1, \dots, u_4, \eta(t)))^T.
 \end{aligned}
 \tag{4}$$

Here,  $(\cdot)^T$  denotes transposition operator.

In the model (1) we assume that  $f \in C^1(\Omega_T)$  is Lipschitz continuous and the function  $\eta(t) \in C(\Omega_T)$  represents the unknown drug efficacy which belongs to the set of admissible functions  $M_\eta$ :

$$M_\eta = \{ \eta(t) : \eta(t) \in [0, 1] \text{ in } \Omega_T, \eta(t) = 0 \text{ outside of } \Omega_T \}.
 \tag{5}$$

The control parameter  $\eta$  is dosage of the reverse transcriptase inhibitor. This parameter protects the cells and prevents infection. In this work we assume that all parameters in system (1) are constants except the control parameter  $\eta$  which depends on time, e.,e.,  $\eta = \eta(t)$ ,  $t \in [0, T]$ . This means that the control parameter  $\eta = \eta(t)$  tells us which dosage of the reverse transcriptase inhibitor should be given to the concrete patient at any time moment  $t$  for  $t \in [0, T]$ . Thus, personalized determination of this parameter for every individual is of vital importance for treatment of HIV.

### Existence and Lyapunov Stability of Steady State

Let us now assume that the parameter  $\eta$  in system (1) is constant. That is,  $\eta(t) \equiv c \in (0, 1)$ . Setting  $\frac{du}{dt}$  in (1) to zero and solving for  $u_1, u_2, u_3$  and  $u_4$  we can see that there are two possible steady states: an infected and an uninfected one [42].

The uninfected steady state is given by

**Table 2** Stable steady states for different values of  $\eta$ , while keeping the other parameters fixed

$\eta$	$u_1$	$u_2$	$u_3$	$u_4$
0.0	116	21	33	3549
0.1	129	23	32	3483
0.2	145	26	31	3402
0.3	166	28	30	3298
0.4	194	32	29	3162
0.5	233	36	27	2975
0.6	291	41	25	2702
0.7	388	45	21	2269
0.8	581	44	14	1469
0.9	1000	0	0	0
1.0	1000	0	0	0

$$\begin{cases} u_1 = \frac{s}{\mu}, \\ u_2 = 0, \\ u_3 = 0, \\ u_4 = 0, \end{cases} \tag{6}$$

and the infected steady state is achieved when

$$\begin{cases} u_1 = \frac{(\mu_1 + \alpha + b)c}{N\alpha k(1-\eta)}, \\ u_2 = \frac{s - \mu u_1}{\mu_1 + \alpha(1-\eta)}, \\ u_3 = \frac{\alpha(1-\eta)u_2}{N\delta u_3}, \\ u_4 = \frac{\delta}{c}. \end{cases} \tag{7}$$

In [42] was shown that the infected steady state can exist only when  $\eta$  is less than the following critical value

$$\eta_{crit} = 1 - \frac{\mu c(\mu_1 + \alpha + b)}{N\alpha k s}. \tag{8}$$

For our system of parameters, presented in Table 1, this critical value is  $\eta_{crit} \approx 0.88375$ . Whenever  $\eta \geq \eta_{crit}$  only the uninfected steady state can exist.

Plugging in the values of Table 1 into (7) for  $\eta < \eta_{crit}$ , or (6) if  $\eta \geq \eta_{crit}$ , we obtain the numerical values for solutions  $(u_1, u_2, u_3, u_4)^T$  of (1) presented in the Table 2.

### Stability of Solutions

Let us define

$$\begin{cases} \mu_m = \min\{\mu, \mu_1\}, \\ \Xi = \frac{s}{\mu_m}, \\ \Phi := \Phi(\eta) = \frac{\alpha s(1-\eta)}{\mu_m}, \\ \Psi := \Psi(\eta) = \frac{N\alpha s(1-\eta)}{\mu_m c}, \end{cases} \tag{9}$$

where  $\mu, \mu_1, s$  etc. are the parameters of (1). Consider the set

$$\Gamma(\eta) = \{(u_1, u_2, u_3, u_4) \in \mathbb{R}^4 : 0 \leq u_1 \leq \Xi, 0 \leq u_2 \leq \Xi, 0 \leq u_3 \leq \Phi, 0 \leq u_4 \leq \Psi\}. \tag{10}$$

It can be proven [42] that if  $u(0) \in \Gamma(\eta)$ , then the solution trajectories of (1) will stay inside  $\Gamma(\eta)$  for all  $t \in \Omega_T$ .

**Remark 1** It is not required that  $\eta$  is constant. As long as  $\eta \in M_\eta$ , we may allow  $\eta(t)$  to vary with time.

For our parameters presented in Table 1, these bounds are quantitatively defined as

$$\begin{cases} 0 \leq u_1 \leq 1000, \\ 0 \leq u_2 \leq 1000, \\ 0 \leq u_3 \leq 1538.5(1 - \eta), \\ 0 \leq u_4 \leq 166667(1 - \eta). \end{cases} \tag{11}$$

(Table 3) shows upper limits for  $u_3$  and  $u_4$  for different values of  $\eta$ .

It can furthermore be proven that if and only if  $\eta \geq \eta_{crit}$  the uninfected state is globally asymptotically Lyapunov stable. On the other hand, if the steady state exists, then it is locally asymptotically Lyapunov stable whenever the following condition is satisfied [42]

$$\Delta C - A^2 D > 0, \tag{12}$$

where

$$\begin{aligned} A &= \mu + ku_4 + \mu_1 + \alpha + b + \delta + c, \\ B &= (c + \delta)(\alpha + \mu_1 + \mu + ku_4 + b) + c\delta + \mu(\mu_1 + \alpha + b) + ku_4(\mu_1 + (1 - \eta)\alpha), \\ C &= c\delta(\mu + ku_4) + (c + \delta)(\mu\mu_1 + \mu\alpha + \mu b + \mu_1ku_4 + (1 - \eta)\alpha ku_4), \\ D &= c\delta ku_4(\mu_1 + \alpha(1 - \eta)), \\ \Delta &= AB - C. \end{aligned} \tag{13}$$

We can calculate that, when  $\eta$  is constant and the other parameter values are chosen as in Table 1, then the infected steady state is locally asymptotically stable for all values of  $\eta$  such that  $\eta$  is less than the critical value,  $\eta_{crit} \approx 0.88$ . Figure 3 illustrates this statement.

Thus, if  $\eta$  is constant, and less than the critical value  $\eta_{crit} \approx 0.88$ , it suffices to know the solution of (1) at steady state to deduce  $\eta$ .

Although it is often a reasonable assumption that the drug efficacy is constant for a given individual, viruses mutate readily, which can alter the efficacy of a RT-inhibitor. Thus, it is interesting to know how to determine  $\eta(t)$  when it is not constant. So let us for the remainder of this note consider the case when  $\eta(t)$  is not necessarily constant.

### Well-Posedness of the Forward Problem

Let  $D = \Omega_T \times \Gamma(\eta)$  be the bounded domain. Let functions  $u(t), f(t, u(t))$  are defined as in (4). Further, let  $f(t, u(t))$  be a continuous function for  $t$  in  $\Omega_T$  and Lipschitz continuous function for  $u(t)$  in  $D$ . Then  $f(t, u(t))$  is clearly Lipschitz continuous on the compact set  $\Gamma(\eta) \times \Omega_T$ . In other words,  $\exists L = const. : \forall t \in \Omega_T, \forall u^1(t), u^2(t) \in D,$

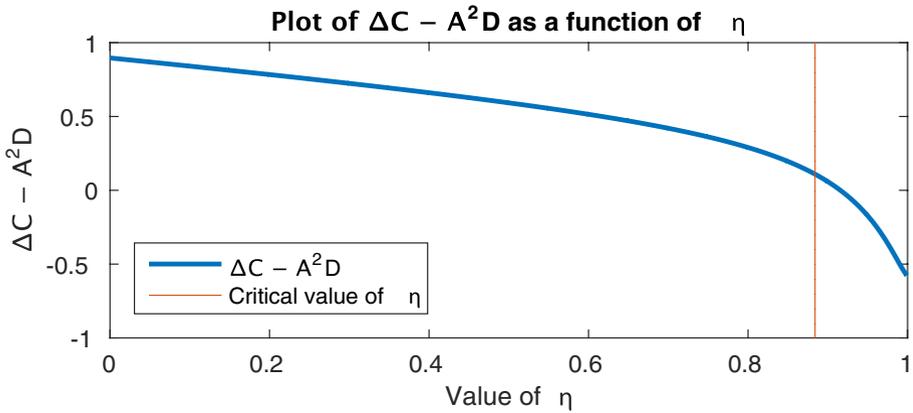


Fig. 3  $\Delta C - A^2D$  plotted as a function of  $\eta$ . Note that  $\Delta C - A^2D > 0 \forall \eta < 0.88$

**Table 3** Upper limit for the positive invariant set  $\Gamma(\eta)$ . The integer parts of fractional numbers is always reported as the upper bound

$\eta$	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$u_3$	1538	1384	1230	1076	923	769	615	461	307	153
$u_4$	166666	150000	133333	116666	100000	83333	66666	50000	33333	16666

$$\|f(t, u^1) - f(t, u^2)\| \leq L\|u^1(t) - u^2(t)\|. \tag{14}$$

Thus, using the Picard–Lindelöf theorem (Theorem 2.2 in [43]) one can prove that, for given initial condition  $u(0)$ , the model problem (1) has a unique solution. Furthermore, the solution depends continuously on data of the problem (1) in the following sense (Theorem 2.8 in [43]):

**Proposition 2** *Let  $u^1(t), u^2(t)$  be two solutions of the problems*

$$\frac{du^i}{dt} = f^i(t, u^i(t)), \quad u^i(t_0) = u^i_0, \quad i = 1, 2$$

*with perturbed initial conditions*

$$\|u^1(0) - u^2(0)\| \leq \delta$$

*and perturbed right hand sides*

$$\|f^1(t, u^1(t)) - f^2(t, u^2(t))\| \leq \epsilon, \quad \forall t \in \Omega_T.$$

Then

$$\|u^1(t) - u^2(t)\| \leq \|u^1(0) - u^2(0)\|e^{Lt} + \frac{\|f^1(t, u^1(t)) - f^2(t, u^2(t))\|}{L}(e^{Lt} - 1), \tag{15}$$

where  $L$  is the Lipschitz constant. If the initial values are equal then on  $\Omega_T = [0, T]$  we have

$$\|u^1(t) - u^2(t)\| \leq \frac{\|f^1(t, u^1(t)) - f^2(t, u^2(t))\|}{L} (e^{Lt} - 1). \tag{16}$$

And since  $f$  is clearly continuous with respect to  $\eta$  it follows that the solution to (1) must be continuous with respect to  $\eta$ .

### Inverse Problems and Ill-Posedness

Since the parameter identification problem can be considered as an inverse ill-posed problem it is clear that this problem is difficult to solve properly. In this section we show how parameter identification problem can be solved accurately via construction of proper Tikhonov regularization functional.

Let us consider the following problem: Let  $B_1$  and  $B_2$  be Banach spaces. Let  $G \subseteq B_1$  be an open set in  $B_1$  and  $F : G \rightarrow B_2$  an operator. Let  $y \in B_2$  be given, and suppose we want to find  $x \in G$  such that

$$F(x) = y. \tag{17}$$

Problems of this sort, when you want to identify  $x$  in (17), given observations,  $y$ , are called *inverse problems*. A special class of inverse problems are called parameter identification problems (PIP), i.e.  $x$  is some parameter of a differential equation, and  $F(x)$  is the solution of the differential equation, with this parameter.

**Definition 1** Problem (17) is said to be well-posed by Hadamard if it satisfies the following conditions [45, 46]:

1. **Existence:** For each  $y \in B_2$  there is an  $x = x(y)$  such that  $F(x) = y$ .
2. **Uniqueness:** For each  $y \in B_2$  there is not more than one  $x = x(y)$  such that  $F(x) = y$ .
3. **Stability:** For each  $y$  such that a unique solution of (17) exists, the solution  $x = x(y)$  is a continuous function of  $y$ .<sup>1</sup>

**Definition 2** Problem (17) is said to be ill-posed if it is not well-posed.

PIP and other inverse problems are often ill-posed. Ill-posedness means that it is difficult to solve (17) numerically, since measurement errors, or even errors induced by finite-precision computer arithmetic, can have disastrous consequences. Let  $y^*$  denote noiseless observations,  $\delta > 0$  be the noise level, and  $B_\delta[y^*] = \{y : \|y - y^*\|_{B_2} \leq \delta\}$ . The solution to the slightly perturbed equation  $F(x) = y_\delta$  (with  $y_\delta \in B_\delta[y^*]$ ) could be entirely different from the solution to  $F(x) = y^*$ . Perhaps a solution to  $F(x) = y_\delta$  does not even exist. No matter how small  $\delta$  is. A generally ill-posed problem (17) can be well-posed if we consider the restriction of  $F$  in (17) to certain subsets of its domain. In this case is introduced the following definition.

<sup>1</sup> We will call a problem *well-conditioned* if it is well-posed and such that a small change in data results in a small change in the solution: even if  $x(y)$  is continuous it might still be very sensitive to changes in  $y$ .

**Conditional well-posedness**

Let  $B_1$  and  $B_2$  be Banach spaces. Suppose  $G \subset B_1$  is the closure of an open subset in  $B_1$ . Let  $F : G \rightarrow B_2$  be a continuous operator. Assume that  $y^* \in F(G)$  is our ideal noiseless data, and pick a noise level  $\delta > 0$ . Suppose we want to solve

$$F(x) = y_\delta, \tag{18}$$

where  $y_\delta \in B_\delta[y^*]$ . This problem is called *conditionally well-posed on  $G$*  if it satisfies the following conditions [45, 46]:

1. **Existence:** It is a priori known<sup>2</sup> that there exists an ideal solution  $x^* = x^*(y^*) \in G$  for an ideal noiseless data  $y^*$ .
2. **Uniqueness:** The operator  $F : G \rightarrow B_2$  is one-to-one.
3. **Stability:** The inverse operator  $F^{-1}$  is continuous on  $F(G)$ .

**Definition 3** The set  $G$  in Definition 3 is called the *correctness set* of the problem (18).

Continuity of the inverse operator  $F^{-1}$  can be guaranteed if the domain of  $F$  is compact. Hence, any compact set with nonempty interior such that  $F$  is one-to-one is a correctness set. This suggests a method to solve (18) by choosing a suitable correctness set,  $G$ , and then finding a  $x \in G$  such that  $\|F(x) - y_\delta\|$  is as small as possible. The Tikhonov’s theorem offers a such method.

**Theorem 3** (Tikhonov [44]) *Let  $B_1$  and  $B_2$  be Banach spaces, and  $U \subset B_1$  a compact set. Let  $F : U \rightarrow B_2$  be a continuous one-to-one operator and  $V = F(U)$ . Then  $F^{-1} : V \rightarrow B_1$  is a continuous operator.*

For a proof of this fundamental theorem see, for example [13, 44].

**Quasi-Solution**

Let  $H_1$  and  $H_2$  be Hilbert spaces,<sup>3</sup> and assume that  $F : G \rightarrow H_2$  is a continuous mapping defined on a compact correctness set,  $G \subset H_1$ . Let  $\delta > 0$  and assume, as before, that we want to solve

$$F(x) = y_\delta, \tag{19}$$

with  $y_\delta$  defined as before. We know that a solution exists for perfect data  $y^*$ , but in general (19) has no solution, since  $y_\delta \notin F(G)$  (implying that we are dealing with an ill-posed problem). Our goal in this, and the following subsection is to sketch how to construct a family of approximate solutions  $\{x_\delta\}$  in  $G$  that converges to  $x^*$  as  $\delta \rightarrow 0$ . Let us define

<sup>2</sup> The rationale behind this is the assumption that the problem to be solved is a model of some natural phenomenon. And since the phenomenon apparently exists, a solution of the equation describing it must also exist. At least if we assume that the natural phenomenon in question is exactly represented by the mathematical model.

<sup>3</sup> We require that the spaces are Hilbert spaces, rather than arbitrary Banach spaces in order for the closest point property to hold. However, it suffices that they are so-called reflexive Banach spaces, see [20].

$$J_{y_\delta}(x) = \|F(x) - y_\delta\|_{B_2}^2. \tag{20}$$

Since  $F$  is continuous, it takes compact sets to compact sets, thus  $F(G)$  is compact in  $B_2$ . And since  $F(G)$  is a compact subset of a Hilbert space, and therefore closed, a minimum of (20) exists (and if  $F(G)$  also happens to be convex, this minimum is unique). Any  $x \in G$ , unique or not, that minimizes  $J_{y_\delta}$  in (20) is called a *quasi-solution* to (19).

Since the inverse mapping,  $F^{-1}$ , is continuous by the Theorem 3, and is defined on a compact metric space, it admits a modulus of continuity,  $\omega_{F^{-1}}$ .<sup>4</sup> From Theorem 1.5 in [13] it follows that, given  $y_\delta \in B_2$ , then for any quasi-solution  $x_\delta \in \min_{x \in G} J_{y_\delta}(x)$  the following error estimate holds:

$$\|x_\delta - x^*\|_{B_1} \leq \omega_{F^{-1}}(2\delta), \tag{22}$$

where  $\omega_{F^{-1}}(z)$  is the modulus of continuity of the inverse operator  $F^{-1}$ . Thus  $x_\delta \rightarrow x^*$  as  $\delta \rightarrow 0$ . Hence, we can take a sequence of quasi-solutions to be our desired family.

However, sometimes the set of all plausible solutions to (19) does not form a compact set, and in these cases,  $F$  need not be continuous on the set of all plausible solutions—such problems are called essentially ill-posed. Thus, it is not obvious how to choose a suitable correctness set. And even if all the plausible solutions form a compact correctness set, the minimum of (20) may not be unique: there may be local minima or regions where the gradient of the functional is very small, where a minimization algorithm could get trapped. In the next subsection, we will discuss how a stable solution to essentially ill-posed problems could be obtained in practice.

### The Tikhonov Functional

The Tikhonov functional makes sure that when minimizing (20), we will stay in the neighborhood of some point,  $x_0$ , which is a priori known to be close to the true solution,  $x^*$ . A general Tikhonov functional is given below

$$J_\gamma(x) = \frac{1}{2} \|F(x) - y\|_{B_2}^2 + \frac{\gamma}{2} \|x - x_0\|_{B_1}^2. \tag{23}$$

The first term is essentially the same as in (20), the second term is the regularization term and  $\gamma := \gamma(\delta)$  is the regularization parameter. The regularization parameter can be chosen as, for instance,

$$\gamma(\delta) = \delta^{2\mu}, \tag{24}$$

where  $\mu \in (0, 1)$ , see details in [4].

In general, the Tikhonov functional (23) might not actually attain its infimum; we can only guarantee the existence of the minimizing sequence,  $\{x_k\}$ . However, without loss of generality, we can assume that  $G$  is the closure of an open and bounded set containing the initial guess,  $x_0$ , the (bounded) minimizing sequence,  $\{x_k\}$ , and the exact solution,  $x^*$ .

<sup>4</sup> A modulus of continuity is a function  $\omega : [0, \infty] \rightarrow [0, \infty]$  such that

$$\lim_{x \rightarrow 0^+} \omega(x) = \omega(0) = 0. \tag{21}$$

The function  $f$  admits  $\omega$  as a modulus of continuity if and only if  $\|f(x) - f(y)\| \leq \omega(\|x - y\|)$ . In particular,  $f$  has a modulus of continuity if and only if it is uniformly continuous.

Hence, if we consider finite dimensional Hilbert spaces, the Tikhonov functional, defined on  $G$ , would have a minimum, since closed and bounded sets on finite dimensional Hilbert spaces are compact, and functionals defined on compact sets attain their infimum according to the Weierstrass' extreme value theorem. Suppose now that  $G$  is convex and that (23) is Fréchet differentiable,<sup>5</sup> with a Fréchet derivative that is uniformly bounded and Lipschitz continuous. Then one can prove, see [13, 14], that for given noise level and regularization parameter (23) is locally strongly convex in a neighborhood of its minimum and that  $x^*$  is also contained in this neighborhood if  $\|x^* - x_0\|$  is small enough.

Assume that we have a single noise  $\delta$  and our goal is to minimize (23). Let  $\gamma(\delta)$  is chosen as in (24), then it can be proven, see [28], that there exists a  $\delta_0$  such that

$$\delta \in (0, \delta_0) \implies \|x_k - x^*\| \leq \xi \|x_0 - x^*\|, \tag{26}$$

in particular it follows that if (23) attains a minimum, any  $x_k \in \min_{x \in G} J(x)$  would be a better approximation to  $x^*$  than  $x_0$ , if the noise level is small enough.

Thus, to sum up, under reasonable assumptions discussed above, a minimum of (23) exists and is a better approximation than the starting guess,  $x_0$ . And under reasonable assumptions, and if the initial guess is good enough, there is only one unique point that is the global minimum, and we do not need to worry about local minima. These facts explain why Tikhonov regularization is so useful for solving ill-posed problems. To find the zero of the Fréchet derivative, one can use common minimization techniques, such as the conjugate gradient method (CGM) or the method of steepest descent. Obviously, the minimum of the Tikhonov functional will not be exactly the same as the quasi-solution if the noise level and regularization parameter are constants. On the other hand, by letting the regularization parameter  $\gamma$  decrease for each iteration of the minimization algorithm we will have a minimum of the Tikhonov functional that approaches the quasi-solution. In [4] it was suggested that  $\gamma$  can be updated as

$$\gamma_k = \frac{\gamma_0}{(k + 1)^p}, \tag{27}$$

where  $p \in (0, 1]$  and  $k = 0, 1, 2, \dots$

From what have been discussed, it is clear that a good first guess is essential for successful identification of the desired parameter. Of course, we do not in general have any idea at all what the solution to (17) might be, and therefore, in general, we need to devise some kind of globally convergent algorithm to solve ill-posed PIP.

## The Parameter Identification Problem

To formulate the parameter identification problem we assume that all parameters in system (1) are known except the control parameter  $\eta(t)$  which describes efficacy of the drug. The typical values of parameters  $\{s, \mu, k, \mu_1, \alpha, b, \delta, c, N\}$  in (1) are taken from [42] and they are described in Table 1.

<sup>5</sup> A continuous linear operator between Banach spaces  $A : B_1 \rightarrow B_2$  is called the Fréchet derivative of the operator  $T : B_1 \rightarrow B_2$  at  $x \in B_1$  if

$$\lim_{\|h\| \rightarrow 0} \frac{\|T(x+h) - T(x) - Ah\|}{\|h\|} = 0. \tag{25}$$

**Parameter Identification Problem (PIP).** Assume that conditions (5) hold and parameters  $\{s, \mu, k, \mu_1, \alpha, b, \delta, c, N\}$  in system (1) are known. Assume further that the function  $\eta(t) \in M_\eta$  is unknown inside the domain  $\Omega_T$ . The PIP is: determine  $\eta(t)$  for  $t \in \Omega_T$ , under the condition that the virus population function  $g(t)$  is known

$$u_4(t) = g(t), \quad t \in [T_1, T_2], \quad 0 \leq T_1 < T_2 \leq T. \tag{28}$$

Here, the function  $g(t)$  presents observations of the function  $u_4(t)$  inside the observation interval  $[T_1, T_2]$ .

Note, that we solve the PIP on the time interval  $[0, T]$  and assume that observations of  $g(t)$  can even be on the more narrow interval  $[T_1, T_2] \subset [0, T]$ . “Numerical Results” section show that reconstruction of the parameter  $\eta(t)$  is not very good on the time interval where observations are not available and thus, observations of the virus population function  $u_4(t)$  should be taken as early as possible from the date when the virus started to be reproduced in the body of the host.

### Optimization Method

Let  $H$  be a Hilbert space of functions defined in  $\Omega_T$ . To determine  $\eta(t)$ ,  $t \in [0, T]$  in PIP we construct the Tikhonov functional (23) in the following form:

$$J(\eta) = \frac{1}{2} \int_{T_1}^{T_2} (u_4(t) - g(t))^2 z_\zeta(t) dt + \frac{1}{2} \gamma \int_0^T (\eta - \eta^0)^2 dt. \tag{29}$$

Here, the solution  $u_4(t)$  of the system (1) with parameter  $\eta(t)$ ,  $g(t)$  is the observed virus population function,  $\eta^0$  is the initial guess for the parameter  $\eta(t)$  and  $\gamma \in (0, 1)$  is the regularization parameter,  $z_\zeta(t)$ ,  $\zeta \in (0, 1)$  is smoothness function which can be defined similarly to [10]. Our goal now is to minimize the Tikhonov functional (29) with respect to the function  $\eta(t) \in H$ .

To find the function  $\eta(t) \in H$  which minimizes the Tikhonov functional (29) we seek for a stationary point of (29) with respect to  $\eta$  which satisfies

$$J'(\eta)(\bar{\eta}) = 0, \quad \forall \bar{\eta} \in H. \tag{30}$$

To find minimum of (29) we use constrained optimization with the standard Lagrangian approach [1, 38] and introduce the following Lagrangian

$$L(v) = J(\eta) + \sum_{i=1}^4 \int_0^T \lambda_i \left( \frac{du_i}{dt} - f_i \right) dt, \tag{31}$$

where  $u(t) = (u_1(t), u_2(t), u_3(t), u_4(t))$  is the solution of the system (1),  $\lambda(t)$  is the vector of Lagrange multipliers  $\lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_4(t))$  and  $v = (\lambda, u, \eta)$ .

Let us introduce following spaces needed for further analysis

$$\begin{aligned}
 H_u^1(\Omega_T) &= \{f \in H^1(\Omega_T) : f(0) = 0\}, \\
 H_\lambda^1(\Omega_T) &= \{f \in H^1(\Omega_T) : f(0) = 0, f(T) = 0\}, \\
 U &= H_u^1(\Omega_T) \times H_\lambda^1(\Omega_T) \times C(\Omega_T),
 \end{aligned}
 \tag{32}$$

where all functions are real valued.

To derive the Fréchet derivative of the Lagrangian (31) we assume that functions  $v = (\lambda, u, \eta)$  can be varied independently of each other in the sense that

$$L'(v)(\bar{v}) = 0, \quad \forall \bar{v} = (\bar{\lambda}, \bar{u}, \bar{\eta}) \in U. \tag{33}$$

Thus, we consider  $L(v + \bar{v}) - L(v)$ , single out the linear part with respect to  $v$  of the obtained expression and neglect all nonlinear terms. The optimality condition (33) means that for all  $\bar{v} \in U$  we have

$$L'(v; \bar{v}) = \frac{\partial L}{\partial \lambda}(v)(\bar{\lambda}) + \frac{\partial L}{\partial u}(v)(\bar{u}) + \frac{\partial L}{\partial \eta}(v)(\bar{\eta}) = 0, \tag{34}$$

i.e., every component of (34) should be zero out. Thus, the optimality conditions (33) yields

$$\begin{aligned}
 0 &= \frac{\partial L}{\partial \lambda}(v)(\bar{\lambda}) = -\alpha \int_0^T u_2(\lambda_1 - \lambda_3)\bar{\eta} dt \\
 &\quad + \int_0^T (u_1 - s + ku_1u_4 + \mu u_1 - (\eta\alpha + b)u_2)\bar{\lambda}_1 dt \\
 &\quad + \int_0^T (u_2 - ku_1u_4 + (\mu_1 + \alpha + b)u_2)\bar{\lambda}_2 dt \\
 &\quad + \int_0^T (u_3 - (1 - \eta)\alpha u_2 + \delta u_3)\bar{\lambda}_3 dt \\
 &\quad + \int_0^T (u_4 - N\delta u_3 + cu_4)\bar{\lambda}_4 dt \quad \forall \bar{\lambda} \in H_u^1(\Omega_T),
 \end{aligned}
 \tag{35}$$

$$\begin{aligned}
 0 &= \frac{\partial L}{\partial u}(v)(\bar{u}) = - \int_0^T (\dot{\lambda}_1 - \lambda_1 k u_4 - \lambda_1 \mu + \lambda_2 k u_4) \bar{u}_1 dt \\
 &\quad - \int_0^T (\dot{\lambda}_2 - \lambda_2(\mu_1 + \alpha + b) + \lambda_1(\eta\alpha + b) + (1 - \eta)\alpha\lambda_3) \bar{u}_2 dt \\
 &\quad - \int_0^T (\dot{\lambda}_3 - \lambda_3\delta + \lambda_4 N\delta) \bar{u}_3 dt \\
 &\quad - \int_0^T (\dot{\lambda}_4 - \lambda_4 c - \lambda_1 k u_1 + \lambda_2 k u_1) \bar{u}_4 dt + \int_{T_1}^{T_2} (u_4 - g) z_\zeta \bar{u}_4 dt \quad \forall \bar{u} \in H_\lambda^1(\Omega_T),
 \end{aligned} \tag{36}$$

$$0 = \frac{\partial L}{\partial \eta}(v)(\bar{\eta}) = \gamma \int_0^T (\eta - \eta^0) \bar{\eta} dt + \alpha \int_0^T u_2 (\lambda_3 - \lambda_1) \bar{\eta} dt \quad \forall \bar{\eta} \in C(\Omega_T). \tag{37}$$

The equation (35) corresponds to the forward problem (1)–(2), the equation (36) — to the following adjoint problem

$$\begin{cases}
 \frac{\partial \lambda_1}{\partial t} = \tilde{f}_1(\lambda(t), \eta(t)) = \lambda_1(t) k u_4(t) + \lambda_1(t) \mu - \lambda_2(t) k u_4(t), \\
 \frac{\partial \lambda_2}{\partial t} = \tilde{f}_2(\lambda(t), \eta(t)) = \lambda_2(t) (\mu_1 + \alpha + b) - \lambda_1(t) (\eta(t) \alpha + b) - (1 - \eta(t)) \alpha \lambda_3(t), \\
 \frac{\partial \lambda_3}{\partial t} = \tilde{f}_3(\lambda(t), \eta(t)) = \lambda_3(t) \delta - \lambda_4(t) N \delta, \\
 \frac{\partial \lambda_4}{\partial t} = \tilde{f}_4(\lambda(t), \eta(t)) = \lambda_4(t) c + \lambda_1(t) k u_1(t) - \lambda_2(t) k u_1(t) + (u_4(t) - g) z_\zeta, \\
 \lambda_i(T) = 0, \quad i = 1, \dots, 4,
 \end{cases} \tag{38}$$

which can be rewritten in the compact form as

$$\begin{cases}
 \frac{\partial \lambda}{\partial t} = \tilde{f}(\lambda(t), \eta(t)), \\
 \lambda_i(T) = 0, \quad i = 1, \dots, 4,
 \end{cases} \tag{39}$$

with

$$\begin{aligned}
 \lambda &= \lambda(t) = (\lambda_1(t), \lambda_2(t), \lambda_3(t), \lambda_4(t))^T, \\
 0 &= (\lambda_1(T), \lambda_2(T), \lambda_3(T), \lambda_4(T))^T, \\
 \frac{d\lambda}{dt} &= \left( \frac{\partial \lambda_1}{\partial t}, \frac{\partial \lambda_2}{\partial t}, \frac{\partial \lambda_3}{\partial t}, \frac{\partial \lambda_4}{\partial t} \right)^T, \\
 \tilde{f}(\lambda(t)) &= (\tilde{f}_1, \tilde{f}_2, \tilde{f}_3, \tilde{f}_4)(\lambda(t), \eta(t))^T.
 \end{aligned} \tag{40}$$

The adjoint system should be solved backwards in time with already known solution  $u(t)$  to the forward problem (1)–(2) and a given measurement function  $g(t)$ .

For the case when  $u$  and  $\lambda$  are exact solutions of the forward (1)–(2) and adjoint (39) problems, respectively, to the known function  $\eta$ , we get from (31) that

$$L(v(\eta)) = J(\eta), \tag{41}$$

and thus the Fréchet derivative of the Tikhonov functional can be written as

$$J'(\eta) := J_\eta(u(\eta), \eta) = \frac{\partial J}{\partial \eta}(u(\eta), \eta) = \frac{\partial L}{\partial \eta}(v(\eta)). \tag{42}$$

Using (37) in (42), we get the following expression for the Fréchet derivative of the Tikhonov functional

$$J'(\eta)(t) = \gamma(\eta - \eta^0)(t) + \alpha u_2(\lambda_3 - \lambda_1)(t) = 0, \tag{43}$$

Thus, to find the unknown parameter  $\eta$  which minimizes the Tikhonov functional (29) we can use the following expression

$$\eta = \frac{1}{\gamma} \alpha u_2(\lambda_1 - \lambda_3) + \eta^0. \tag{44}$$

### Finite Element Discretization

For solution of (33) we will use the finite element discretization and consider a partition  $\mathcal{J}_\tau = \{J\}$  of the time domain  $\Omega_T = [0, T]$  into time subintervals  $J = (t_{k-1}, t_k]$  of the time step  $\tau_k = t_k - t_{k-1}$ . We define also the piecewise-constant time-mesh function  $\tau$  such that

$$\tau(t) = \tau_k, \quad \forall J \in \mathcal{J}_\tau. \tag{45}$$

For discretization of the state and adjoint problems we define the finite element spaces  $W_\tau^u \subset H_u^1(\Omega_T)$  and  $W_\tau^\lambda \subset H_\lambda^1(\Omega_T)$  for  $u$  and  $\lambda$ , respectively, as

$$\begin{aligned} W_\tau^u &= \{f \in H_u^1 : f|_J \in P^1(J) \quad \forall J \in \mathcal{J}_\tau\}, \\ W_\tau^\lambda &= \{f \in H_\lambda^1 : f|_J \in P^1(J) \quad \forall J \in \mathcal{J}_\tau\}. \end{aligned} \tag{46}$$

For the function  $\eta(t)$  we also introduce the finite element space  $W_\tau^\eta \subset L_2(\Omega_T)$  consisting of piecewise constant functions

$$W_\tau^\eta = \{f \in L_2(\Omega_T) : f|_J \in P^0(J) \quad \forall J \in \mathcal{J}_\tau\}. \tag{47}$$

We use different finite element spaces since we are working in a finite dimensional space and all norms in finite dimensional spaces are equivalent. Next we denote  $U_\tau = W_\tau^u \times W_\tau^\lambda \times W_\tau^\eta$  such that  $U_\tau \subset U$ .

Now the finite element method for (33) is: find  $v_\tau \in U_\tau$  such that

$$L'(v_\tau; \bar{v}) = 0, \quad \forall \bar{v} \in U_\tau. \tag{48}$$

Since the forward (1)–(2) and adjoint (36) problems are nonlinear their solutions can be found by Newton’s method. For the discretization

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

the variational formulation of the forward problem (1)–(2) for all  $\bar{u} \in H_u^1(\Omega_T)$  is:

$$(u^{k+1}, \bar{u}) - (u^k, \bar{u}) - (\tau_k f(u^{k+1}), \bar{u}) = 0. \tag{49}$$

The finite element method for (1)–(2) will be: find  $u_\tau^{k+1} \in H_u^1(\Omega_T)$  such that for all  $\bar{u} \in H_u^1(\Omega_T)$

$$(u_\tau^{k+1}, \bar{u}) - (u_\tau^k, \bar{u}) - (\tau_k f(u_\tau^{k+1}), \bar{u}) = 0. \tag{50}$$

Denoting

$$\begin{aligned} \tilde{u} &= u_\tau^{k+1}, \\ V(\tilde{u}) &= \tilde{u} - \tau_k f(\tilde{u}) - u_\tau^k \end{aligned} \tag{51}$$

we can rewrite (50) as

$$(V(\tilde{u}), \bar{u}) = 0. \tag{52}$$

For solution  $V(\tilde{u}) = 0$  the Newton’s method can be used for the iterations  $n = 1, 2, \dots$  [17]

$$\tilde{u}^{n+1} = \tilde{u}^n - [V'(\tilde{u}^n)]^{-1} \cdot V(\tilde{u}^n). \tag{53}$$

Here, we can compute the Jacobian  $V'(\tilde{u}^n)$  via definition of  $V(\tilde{u})$  in (51) as

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

where  $I$  is the identity matrix,  $f'(\tilde{u}^n)$  is the Jacobian of  $f$  (the right hand side of the forward problem (1)) at  $\tilde{u}^n$  and  $n$  is the iteration number in Newton’s method. The explicit entries in the Jacobian  $f'(\tilde{u}^n)$  for system (1) are computed as

$$f'(\tilde{u}^n) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} & \frac{\partial f_1}{\partial u_3} & \frac{\partial f_1}{\partial u_4} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} & \frac{\partial f_2}{\partial u_3} & \frac{\partial f_2}{\partial u_4} \\ \frac{\partial f_3}{\partial u_1} & \frac{\partial f_3}{\partial u_2} & \frac{\partial f_3}{\partial u_3} & \frac{\partial f_3}{\partial u_4} \\ \frac{\partial f_4}{\partial u_1} & \frac{\partial f_4}{\partial u_2} & \frac{\partial f_4}{\partial u_3} & \frac{\partial f_4}{\partial u_4} \end{bmatrix} (\tilde{u}^n) = \begin{bmatrix} -ku_{\tau_4}^n - \mu & (\eta\alpha + b) & 0 & -ku_{\tau_1}^n \\ ku_{\tau_4}^n & -(\mu_1 + \alpha + b) & 0 & ku_{\tau_1}^n \\ 0 & (1 - \eta)\alpha & -\delta & 0 \\ 0 & 0 & N\delta & -c \end{bmatrix}.$$

We note that the finite element method (48) will work even in this case, see details in [24].

In a similar way the Newton’s method can be derived for the solution the adjoint problem (39). Since we solve the adjoint problem backwards in time starting from the known  $\lambda(T) = 0$ , we discretize time derivative as

$$\frac{\partial \lambda}{\partial t} = \frac{\lambda^{k+1} - \lambda^k}{\tau_k} \tag{54}$$

for the already known  $\lambda^{k+1}$ , and write the variational formulation of the adjoint problem for all  $\bar{\lambda} \in H_\lambda^1(\Omega_T)$  as

$$-(-\lambda^{k+1} + \lambda^k + \tau_k \tilde{f}(\lambda^k), \bar{\lambda}) = 0. \tag{55}$$

The finite element method for (39) will be: find  $\lambda_\tau^k \in H_\lambda^1(\Omega_T)$  such that for all  $\bar{\lambda} \in H_\lambda^1(\Omega_T)$

$$(\lambda_\tau^k - \lambda_\tau^{k+1} + \tau_k \tilde{f}(\lambda_\tau^k), \bar{\lambda}) = 0. \tag{56}$$

Denoting

$$\begin{aligned} \tilde{\lambda} &= \lambda_{\tau}^k, \\ \tilde{V}(\tilde{\lambda}) &= \tilde{\lambda} + \tau_k \tilde{f}(\tilde{\lambda}) - \lambda_{\tau}^{k+1}, \end{aligned} \tag{57}$$

we can rewrite (56) for all  $\tilde{\lambda} \in H_{\lambda}^1(\Omega_T)$  as

$$(\tilde{V}(\tilde{\lambda}), \tilde{\lambda}) = 0. \tag{58}$$

For solution  $\tilde{V}(\tilde{\lambda}) = 0$  we use again Newton’s method for iterations  $n = 1, 2, \dots$

$$\tilde{\lambda}^{n+1} = \tilde{\lambda}^n - [\tilde{V}'(\tilde{\lambda}^n)]^{-1} \cdot \tilde{V}(\tilde{\lambda}^n). \tag{59}$$

We compute  $\tilde{V}'(\tilde{\lambda}^n)$  using the definition of  $\tilde{V}(\tilde{\lambda})$  in (57) as

$$\tilde{V}'(\tilde{\lambda}^n) = I + \tau_k \tilde{f}'(\tilde{\lambda}^n),$$

where  $I$  is the identity matrix,  $\tilde{f}'(\tilde{\lambda}^n)$  is the Jacobian of  $\tilde{f}$  (the right hand side of the adjoint problem (39)) at  $\tilde{\lambda}^n$ , and  $n$  is the iteration number in Newton’s method. The explicit entries in the Jacobian  $\tilde{f}'(\tilde{\lambda}^n)$  for the adjoint system (39) are given by

$$\tilde{f}'(\tilde{\lambda}^n) = \begin{bmatrix} \frac{\partial \tilde{f}_1}{\partial \lambda_1} & \frac{\partial \tilde{f}_1}{\partial \lambda_2} & \frac{\partial \tilde{f}_1}{\partial \lambda_3} & \frac{\partial \tilde{f}_1}{\partial \lambda_4} \\ \frac{\partial \tilde{f}_2}{\partial \lambda_1} & \frac{\partial \tilde{f}_2}{\partial \lambda_2} & \frac{\partial \tilde{f}_2}{\partial \lambda_3} & \frac{\partial \tilde{f}_2}{\partial \lambda_4} \\ \frac{\partial \tilde{f}_3}{\partial \lambda_1} & \frac{\partial \tilde{f}_3}{\partial \lambda_2} & \frac{\partial \tilde{f}_3}{\partial \lambda_3} & \frac{\partial \tilde{f}_3}{\partial \lambda_4} \\ \frac{\partial \tilde{f}_4}{\partial \lambda_1} & \frac{\partial \tilde{f}_4}{\partial \lambda_2} & \frac{\partial \tilde{f}_4}{\partial \lambda_3} & \frac{\partial \tilde{f}_4}{\partial \lambda_4} \end{bmatrix} (\tilde{\lambda}^n) = \begin{bmatrix} ku_{\tau_4}^n + \mu & -ku_{\tau_4}^n & 0 & 0 \\ -(\eta\alpha + b) & \mu_1 + \alpha + b & (\eta - 1)\alpha & 0 \\ 0 & 0 & \delta & -N\delta \\ ku_{\tau_1}^n & -ku_{\tau_1}^n & 0 & c \end{bmatrix}.$$

Taking into account values of parameters given in Table 1, we observe that  $\det \tilde{f}'(\tilde{\lambda}^n) \neq 0$  as well as  $\det \tilde{V}'(\tilde{\lambda}^n) \neq 0$ . Thus, schemes (53), (59) will converge given the appropriate starting values  $\tilde{u}^1$  and  $\tilde{\lambda}^1$ , correspondingly. For study of convergence of iterative methods we refer to [4].

### A Posteriori Error Estimates

We consider the function  $\eta \in C(\Omega_T)$  as a minimizer of the Lagrangian (31), and  $\eta_{\tau} \in W_{\tau}^n$  its finite element approximation. Let us assume that we know good approximation to the exact solution  $\eta^* \in C(\Omega_T)$ . Let  $g^*(t)$  be the exact data and the function  $g_{\sigma}(t)$  represents the error level in these data. We assume that measurements  $g(t)$  in (28) are given with some noise level (small)  $\sigma$  such that

$$g(t) = g^*(t) + g_{\sigma}(t); \quad g^*, g_{\sigma} \in L_2(\Omega_T), \quad \|g_{\sigma}\|_{L_2(\Omega_T)} \leq \sigma. \tag{60}$$

Accordingly [14] we assume that

$$\gamma = \gamma(\sigma) = \sigma^{2\mu}, \quad \mu \in (0, 1/4), \quad \sigma \in (0, 1) \tag{61}$$

and

$$\|\eta_0 - \eta^*\| \leq \frac{\sigma^{3\mu}}{3}, \tag{62}$$

where  $\eta^*$  is the exact solution of PIP with the exact data  $g^*(t)$ . Let

$$V_\varepsilon(\eta) = \{x \in C(\Omega_T) : \|\eta - x\| < \varepsilon \quad \forall \eta \in C(\Omega_T)\}. \tag{63}$$

Assume that for all  $\eta \in V_1(\eta^*)$  the operator

$$F(\eta) = \frac{1}{2} \int_{T_1}^{T_2} (u_4(\eta, t) - g(t))^2 z_\zeta(t) dt \tag{64}$$

has the Fréchet derivative  $F'(\eta)$  which is bounded and Lipschitz continuous in  $V_1(\eta^*)$  for  $D_1, D_2 = const. > 0$

$$\begin{aligned} \|F'(\eta)\| &\leq D_1 \quad \forall \eta \in V_1(\eta^*), \\ \|F'(\eta_1) - F'(\eta_2)\| &\leq D_2 \|\eta_1 - \eta_2\| \quad \forall \eta_1, \eta_2 \in V_1(\eta^*). \end{aligned} \tag{65}$$

### An A Posteriori Error Estimate for the Tikhonov Functional

In the Theorem 1 we derive an a posteriori error estimate for the error in the Tikhonov functional (29) on the finite element time partition  $\mathcal{J}_\tau$ .

**Theorem 1** *We assume that there exists minimizer  $\eta \in C(\Omega_T)$  of the functional  $J(\eta)$  defined by (29). We assume also that there exists finite element approximation of a minimizer  $\eta_\tau \in W_\tau^\eta$  of  $J(\eta)$ . Then the following approximate a posteriori error estimate for the error  $e = \|J(\eta) - J(\eta_\tau)\|_{L^2(\Omega_T)}$  in the Tikhonov functional (29) holds true*

$$e = \|J(\eta) - J(\eta_\tau)\|_{L^2(\Omega_T)} \leq C_I C \|J'(\eta_\tau)\|_{L^2(\Omega_T)} \|\tau \eta_\tau\|_{L_2(\Omega_T)} \tag{66}$$

with positive constants  $C_I, C > 0$  and where

$$J'(\eta_\tau) = \gamma(\eta_\tau - \eta^0) - \alpha u_{2\tau}(\lambda_{1\tau} - \lambda_{3\tau}). \tag{67}$$

**Proof** We use the definition of the Fréchet derivative to get

$$J(\eta) - J(\eta_\tau) = J'(\eta_\tau)(\eta - \eta_\tau) + R(\eta, \eta_\tau), \tag{68}$$

where  $R(\eta, \eta_\tau) = O((\eta - \eta_\tau)^2)$ ,  $(\eta - \eta_\tau) \rightarrow 0 \quad \forall \eta, \eta_\tau \in W_\tau^\eta$ . The term  $R(\eta, \eta_\tau)$  is small because of assumption (62): we assume that  $\eta_\tau$  is the minimizer of the Tikhonov functional on the mesh  $\mathcal{J}_\tau$  and this minimizer is located in a small neighborhood of the regularized solution  $\eta$ . Because of that we neglect  $R$  in (68). Next, we use the splitting

$$\eta - \eta_\tau = \eta - \eta_\tau^I + \eta_\tau^I - \eta_\tau \tag{69}$$

for  $\eta - \eta_\tau$  in (68) together with Galerkin orthogonality

$$J'(\eta_\tau)(\eta_\tau^I - \eta_\tau) = 0, \quad \forall \eta_\tau^I, \eta_\tau \in W_\tau^\eta \tag{70}$$

to get

$$J(\eta) - J(\eta_\tau) \leq J'(\eta_\tau)(\eta - \eta'_\tau). \tag{71}$$

Here,  $\eta'_\tau$  is a standard interpolant of  $\eta$  on the mesh  $\mathcal{J}_\tau$  [23]. Taking norms in (71), we obtain

$$\|J(\eta) - J(\eta_\tau)\|_{L^2(\Omega_T)} \leq \|J'(\eta_\tau)\|_{L^2(\Omega_T)} \|\eta - \eta'_\tau\|_{L^2(\Omega_T)}, \tag{72}$$

where the term  $\|\eta - \eta'_\tau\|_{L^2(\Omega_T)}$  can be estimated via the interpolation estimate with the constant  $C_I$

$$\|\eta - \eta'_\tau\|_{L^2(\Omega_T)} \leq C_I \|\tau \eta\|_{H^1(\Omega_T)}. \tag{73}$$

We can estimate  $\|\tau \eta\|_{H^1(\Omega_T)}$  in (73) as

$$\begin{aligned} \|\tau \eta\|_{H^1(\Omega_T)} &\leq \sum_J \|\tau_k \eta\|_{H^1(J)} = \sum_J \left\| \left( \eta + \frac{\partial \eta}{\partial t} \right) \tau_k \right\|_{L_2(J)} \\ &\leq \sum_J \left( \|\eta_\tau \tau_k\|_{L_2(J)} + \left\| \frac{[\eta_\tau]}{\tau_k} \tau_k \right\|_{L_2(J)} \right) \\ &\leq \|\tau \eta_\tau\|_{L_2(\Omega_T)} + \sum_J \|[\eta_\tau]\|_{L_2(J)}. \end{aligned} \tag{74}$$

Here,  $[\eta_\tau]$  denote the jump of the function  $\eta_\tau$  over the time intervals  $[t_{k-1}, t_k]$  and  $[t_k, t_{k+1}]$  defined as

$$[\eta_\tau] = \eta_\tau^+ - \eta_\tau^-$$

with functions  $\eta_\tau^-, \eta_\tau^+$  computed on  $[t_{k-1}, t_k]$  and  $[t_k, t_{k+1}]$ , respectively.

Now we substitute above estimate into (72) to get

$$\|J(\eta) - J(\eta_\tau)\|_{L^2(\Omega_T)} \leq C_I \|J'(\eta_\tau)\|_{L^2(\Omega_T)} \left( \|\tau \eta_\tau\|_{L_2(\Omega_T)} + \sum_J \|[\eta_\tau]\|_{L_2(J)} \right) \quad \forall \eta_\tau \in W_\tau^\eta. \tag{75}$$

In the case when  $\eta_\tau \in W_\tau^\eta$  terms with jumps in time disappear and we have a posteriori error estimate

$$\|J(\eta) - J(\eta_\tau)\|_{L^2(\Omega_T)} \leq C_I \|J'(\eta_\tau)\|_{L^2(\Omega_T)} \|\tau \eta_\tau\|_{L_2(\Omega_T)} \quad \forall \eta_\tau \in W_\tau^\eta. \tag{76}$$

□

### A Posteriori Error Estimate of the Minimizer on Refined Meshes

Theorems 2 and 3 present two a posteriori error estimates for a minimizer  $\eta$  of the functional (29). Proof of the next theorem follows from the proof of Theorem 5.1 of [29].

**Theorem 2** *Let  $\eta_\tau \in W_\tau^\eta$  be a finite element approximation on the finite element mesh  $\mathcal{J}_\tau$  of the minimizer  $\eta \in L^2(\Omega_T)$  of the functional (29) with the mesh function  $\tau(t)$ . Then there exists a Lipschitz constant  $D = \text{const.} > 0$  defined by*

$$\|J'(\eta_1) - J'(\eta_2)\| \leq D\|\eta_1 - \eta_2\|, \forall \eta_1, \eta_2 \in L_2(\Omega_T), \tag{77}$$

and interpolation constant  $C_I$  independent on  $\tau$  such that the following a posteriori error estimate for the minimizer  $\eta$  holds true

$$\|\eta_\tau - \eta\|_{L_2(\Omega_T)} \leq \frac{D}{\gamma} C_I \|\tau \eta_\tau\|_{L_2(\Omega_T)} \forall \eta_\tau \in W_\tau^\eta. \tag{78}$$

**Proof** Let  $\eta_\tau$  be the minimizer of the Tikhonov functional (29). The existence and uniqueness of this minimizer is guaranteed by conditions (62) and follows from Theorem 1.9.1.2 of [13]. By this theorem, the functional (29) is strongly convex on the space  $L_2(\Omega_T)$  with the strong convexity constant  $\gamma$ . This implies that

$$\gamma \|\eta_\tau - \eta\|_{L_2(\Omega_T)}^2 \leq |(J'(\eta_\tau) - J'(\eta), \eta_\tau - \eta)|. \tag{79}$$

Here,  $J'(\eta_\tau), J'(\eta)$  are the Fréchet derivatives of the functional (29) given by (43) for respective  $\eta$ .

Since  $\eta$  is the minimizer of the Tikhonov functional (29) then

$$(J'(\eta), \eta) = 0, \quad \forall \eta \in L_2(\Omega_T).$$

Using the splitting

$$\eta_\tau - \eta = (\eta_\tau - \eta_\tau^I) + (\eta_\tau^I - \eta), \tag{80}$$

where  $\eta_\tau^I$  is an interpolant of  $\eta$ , together with the Galerkin orthogonality principle for all  $\eta_\tau, \eta_\tau^I \in W_\tau^\eta$

$$(J'(\eta_\tau) - J'(\eta), \eta_\tau - \eta_\tau^I) = 0 \tag{81}$$

in (79) we obtain

$$\gamma \|\eta_\tau - \eta\|_{L_2(\Omega_T)}^2 \leq |(J'(\eta_\tau) - J'(\eta), \eta_\tau^I - \eta)|. \tag{82}$$

We can estimate the right hand side of (82) using (77) as

$$|(J'(\eta_\tau) - J'(\eta), \eta_\tau^I - \eta)| \leq D\|\eta_\tau - \eta\|_{L_2(\Omega_T)} \|\eta_\tau^I - \eta\|_{L_2(\Omega_T)}.$$

Substituting above equation into (82) we obtain

$$\|\eta_\tau - \eta\|_{L_2(\Omega_T)} \leq \frac{D}{\gamma} \|\eta_\tau^I - \eta\|_{L_2(\Omega_T)}. \tag{83}$$

Using the interpolation property

$$\|\eta_\tau^I - \eta\|_{L_2(\Omega_T)} \leq C_I \|\tau \eta\|_{H^1(\Omega_T)} \tag{84}$$

we obtain a posteriori error estimate for the regularized solution with the interpolation constant  $C_j$ :

$$\|\eta_\tau - \eta\|_{L_2(\Omega_T)} \leq \frac{D}{\gamma} \|\eta_\tau^I - \eta\|_{L_2(\Omega_T)} \leq \frac{D}{\gamma} C_I \|\tau \eta\|_{H^1(\Omega_T)}. \tag{85}$$

We can estimate  $\|\tau \eta\|_{H^1(\Omega_\tau)}$  in (85) similar to (74). Substituting this estimate into the right hand side of (85) we get

$$\|\eta_\tau - \eta\|_{L_2(\Omega_\tau)} \leq \frac{D}{\gamma} C_I \left( \|\tau \eta_\tau\|_{L_2(\Omega_\tau)} + \|[ \eta_\tau ]\|_{L_2(\Omega_\tau)} \right) \forall \eta_\tau \in W_\tau^\eta.$$

In the case when  $\eta_\tau \in W_\tau^\eta$  terms with jumps in time  $[ \eta_\tau ]$  disappear and we have a posteriori error estimate

$$\|\eta_\tau - \eta\|_{L_2(\Omega_\tau)} \leq \frac{D}{\gamma} C_I \|\tau \eta_\tau\|_{L_2(\Omega_\tau)}.$$

□

**Theorem 3** Let  $\eta_\tau \in W_\tau^\eta$  be a finite element approximation on the finite element mesh  $J_\tau$  of the minimizer  $\eta \in L^2(\Omega_T)$  of the functional (29) with the mesh function  $\tau(t)$ . Then there exists an interpolation constant  $C_I$  independent on  $\tau$  such that the following a posteriori error estimate for the minimizer  $\eta$  and the regularization parameter  $\gamma \neq 0$  holds

$$\|\eta_\tau - \eta\|_{L_2(\Omega_\tau)} \leq \sqrt{\frac{\|R(\eta_\tau)\|}{\gamma}} C_I \|\tau \eta_\tau\|_{L_2(\Omega_\tau)} \forall \eta_\tau \in W_\tau^\eta, \tag{86}$$

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

$$R(\eta_\tau)(t) = \gamma(\eta_\tau - \eta^0)(t) + \alpha u_{2\tau}((\lambda_{3\tau} - \lambda_{1\tau})(t)). \tag{87}$$

**Proof** Let again  $\eta_\tau$  be the minimizer of the Tikhonov functional (29). Strong convexity of the functional (29) on the space  $L_2(\Omega_T)$  implies that

$$\gamma \|\eta_\tau - \eta\|_{L_2(\Omega_\tau)}^2 \leq |(J'(\eta_\tau) - J'(\eta), \eta_\tau - \eta)|. \tag{88}$$

Applying splitting (80) to (88) we obtain (82) where the term  $J'(\eta_\tau)$  can be estimated via (43). More precisely, when  $u(t), \lambda(t)$  are exact functions, we have for  $\eta_\tau$ :

$$L(v(\eta_\tau)) = J(\eta_\tau),$$

and thus, for exact functions  $u(t), \lambda(t)$  one can write

$$J'(\eta_\tau) = L'(\eta_\tau) = \gamma(\eta_\tau - \eta^0)(t) + \alpha u_{2\tau}((\lambda_{3\tau} - \lambda_{1\tau})(t)). \tag{89}$$

From (88) and (89) (noting that  $J'(\eta) = 0$ ) we get

$$\|\eta_\tau - \eta\|_{L_2(\Omega_\tau)} \leq \sqrt{\frac{\|R(\eta_\tau)\|}{\gamma}} \|\eta_\tau' - \eta\|_{L_2(\Omega_\tau)}, \tag{90}$$

where  $R(\eta_\tau)$  is the residual defined as in (87).

Using the interpolation property (84) and further the estimate (74) we obtain following a posteriori error estimate for the regularized solution with the interpolation constant  $C_I$ :

$$\|\eta_\tau - \eta\|_{L_2(\Omega_\tau)} \leq \sqrt{\frac{\|R(\eta_\tau)\|}{\gamma}} C_I \left( \|\tau \eta_\tau\|_{L_2(\Omega_\tau)} + \|[ \eta_\tau ]\|_{L_2(\Omega_\tau)} \right) \forall \eta_\tau \in W_\tau^\eta.$$

and since  $\eta_\tau \in W_\tau^\eta$  the above estimate reduces to

$$\|\eta_\tau - \eta\|_{L_2(\Omega_\tau)} \leq \sqrt{\frac{\|R(\eta_\tau)\|}{\gamma}} C_I \|\tau \eta_\tau\|_{L_2(\Omega_\tau)} \quad \forall \eta_\tau \in W_\tau^\eta.$$

□

### Algorithms for Solution of PIP

Here we present two algorithms for solution of PIP:

- CGA—usual conjugate gradient algorithm on a coarse time partition,
- ACGA—time-adaptive conjugate gradient algorithm which minimized the Tikhonov functional (29) on a locally refined meshes in time.

We denote the nodal value of the gradient at the observation points  $\{t_i\}$  by  $G^m(t_i)$  and compute it accordingly to (43) as

$$G^m(t_i) = \gamma(\eta_\tau^m(t_i) - \eta_\tau^0(t_i)) + \alpha u_{2\tau}^m(t_i)(\lambda_{3\tau}^m(t_i) - \lambda_{1\tau}^m(t_i)). \tag{91}$$

The approximate computed solutions  $u_{2\tau}^m$  and  $\lambda_{1,3\tau}^m$  are obtained computationally by Newton’s method with  $\eta := \eta_\tau^m$ . A sequence  $\{\eta_\tau^m\}_{m=1,\dots,M}$  of approximations to  $\eta$  is computed as follows

$$\eta_\tau^{m+1}(t_i) = \eta_\tau^m(t_i) + r^m d^m(t_i), \tag{92}$$

with

$$d^m(t_i) = -G^m(t_i) + \beta^m d^{m-1}(t_i),$$

and

$$\beta^m = \frac{\|G^m(t_i)\|^2}{\|G^{m-1}(t_i)\|^2},$$

where  $d^0(t_i) = -G^0(t_i)$  and  $G^m(t_i)$  is the gradient vector which is computed by (91) in time moments  $t_i$ . In (92) the parameter  $r^m$  is the step-size in the gradient update at the iteration  $m$  which is computed as

$$r^m = -\frac{(G^m, d^m)}{\gamma \|d^m\|^2}. \tag{93}$$

**Algorithm 1** Conjugate Gradient Algorithm (CGA).

- 1: Choose time partition  $\mathcal{J}_\tau$  of the time interval  $(0, T)$ . Start with the initial approximations  $\eta_\tau^0$  and compute the sequence of  $\eta_\tau^m$  for all  $m > 0$  in the following steps.
- 2: Compute solutions  $u_\tau^m = u_\tau(t, \eta_\tau^m)$ ,  $\lambda_\tau^m = \lambda_\tau(t, \eta_\tau^m)$  of the state (1) and adjoint (39) problems via (53), (59), respectively, using Newton’s method on the time partition  $\mathcal{J}_\tau$ .
- 3: Compute gradient  $G^m(t_i)$  on the time partition  $I_\tau$  by (91).
- 4: Update the unknown parameter  $\eta := \eta_\tau^{m+1}$  using (92) as

$$\eta_\tau^{m+1}(t_i) = \eta_\tau^m(t_i) + r^m d^m(t_i).$$

- 5: Compute residual  $R(\eta_\tau^m)$  using (87) with solutions  $u_\tau(t, \eta_\tau^m)$ ,  $\lambda_\tau(t, \eta_\tau^m)$  of the state (1) and adjoint (39) problems.
- 6: For the tolerance  $0 < \theta < 1$  chosen by the user, stop computing the functions  $\eta_\tau^m$  if either  $\|R(\eta_\tau^m)\|_{L_2(\Omega_T)} \leq \theta$ , or norms of residuals  $\|R(\eta_\tau^m)\|_{L_2(\Omega_T)}$  abruptly grow, or norms of computed  $\|\eta_\tau^m\|_{L_2(\Omega_T)}$  are stabilized. Otherwise, set  $m := m + 1$  and go to Step 2.

In the adaptive algorithm ACGA we have used Theorem 3 for the error  $e = \|\eta_\tau - \eta\|_{L_2(\Omega_T)}$  on locally refined meshes. More precisely, first we choose tolerance  $0 < \theta < 1$  and run adaptive algorithm until

$$e = \|\eta_\tau - \eta\|_{L_2(\Omega_T)} \leq \theta.$$

For the time-mesh refinements we propose following refinement procedure based on the Theorem 3.

**The Time Mesh Refinements Criterion**

Refine the time-mesh  $\mathcal{J}_\tau$  in neighborhoods of those time-mesh points  $t \in \Omega_T$  where the residual  $|R(\eta_\tau)(t)|$  defined in (87) attains its maximal values. More precisely, let  $\beta_1 \in (0, 1)$  be the tolerance number. Refine the time-mesh in such subdomains of  $\Omega_T$  where

$$|R(\eta_\tau)(t)| \geq \beta_1 \max_{\Omega_T} |R(\eta_\tau)(t)|.$$

Using the above mesh refinement recommendation we propose the following time-adaptive algorithm in computations:

**Algorithm 2** Adaptive Conjugate Gradient Algorithm (ACGA).

- 1: Perform steps 1-6 in CGA algorithm. Let  $M$  be the final iteration in CGA algorithm.
- 2: Refine the time mesh  $\mathcal{J}_\tau$  at all points where

$$|R(\eta_\tau^M)(t)| \geq \beta_1 \max_{\Omega_T} |R(\eta_\tau^M)(t)|. \tag{94}$$

Here the tolerance number  $\beta_1 \in (0, 1)$  is chosen by the user.

- 3: Construct a new time partition  $\mathcal{J}_\tau$  of the time interval  $(0, T)$ . Interpolate the initial approximation  $\eta_0$  from the previous time partition to the new time partition. Next, perform steps 1-6 in CGA algorithm on the new time partition.
- 4: Stop time partition refinements if norms of residuals  $\|R(\eta_\tau^M)\|_{L_2(\Omega_T)}$  either increase or stabilize, compared with the previous time partition.

## Numerical Results

In this section we present several numerical results which show performance and effectiveness of the time-adaptive reconstruction of unknown parameter  $\eta(t), t \in [0, T]$  in PIP using ACGA algorithm. Numerical tests are performed in Matlab R2019b using the developed code for solution of the studied problem available for download at [31]. Numerical results of reconstruction of function  $\eta(t)$  using usual conjugate gradient Algorithm 1 on the nonrefined time-meshes are presented in [25]. We note that observations of all  $u_i, i = 1, 2, 3, 4$  functions in system (1) were used in [25].

The goal of numerical tests of this note is to determine the unknown function  $\eta(t)$  from observation of the virus population function  $u_4(t)$  in (1) on the interval  $[T_1, T_2] \subset [0, T], 0 \leq T_1 < T_2 \leq T$ . In all numerical tests assumed that parameter  $\eta(t)$  satisfy conditions (5) and is unknown in the system (1), but all other parameters  $\{s, \mu, k, \mu_1, \alpha, b, \delta, c, N\}$  of this system are known and their values are chosen as in the Table 1. The observation interval  $[T_1, T_2]$  is such that  $T_2 = T = 300$ , but  $T_1$  is taken differently in different tests since observations of the virus population function  $u_4(t)$  can be taken after the first 3 – 9 weeks since the virus started to be reproduced in the body of host.

For generation of data  $u_4(t) = g(t)$  the problem (1)–(2) was solved numerically with exact values of the test model function  $\eta(t)$ . For solution of problem (1)–(2) was used Newton’s method presented in “The Parameter Identification Problem”. Next, the random noise was added to the observed solution  $u_4(t)$  as

$$u_{4\sigma}(t) = u_{4\sigma}(t)(1 + \sigma\alpha), \tag{95}$$

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

In Algorithms 1, 2 it is of vital importance to take initial guess  $\eta^0$  such that it satisfy condition (62) which means that  $\eta^0$  is located in the close neighborhood of the exact solution. This condition is fulfilled in our PIP since we can compute explicitly values of the parameter  $\eta(t)$  on the initial non-refined time mesh using, for example, the third equation of system (1) as

$$\eta(t) = 1 - \frac{\frac{\partial u_3(t)}{\partial t} + \delta u_3(t)}{\alpha u_2(t)}. \tag{96}$$

We used following discretised version of this equation to get initial guess  $\eta_\tau^0$

**Table 4** Test 1. Relative errors  $e_\eta$  computed for reconstruction of the function  $\eta(t) = 0.7e^{-t} + 0.05, t \in [0, 300]$  for  $T_1 = 25, 50, 100$  on different locally adaptively refined time-meshes

$\sigma$ nr. of ref.	5 %	10%	20%	40%
$T_1 = 25$				
0	0.1893	0.2022	0.2129	0.2203
1	0.1151	0.1223	0.1279	0.2008
2	0.0470	0.0391		
3	0.0354			
4	0.0242			
$T_1 = 50$				
0	0.1917	0.1933	0.1639	0.3498
1	0.1194	0.1267	0.1027	0.2990
2	0.0684	0.0550	0.1002	0.1755
3	0.0337	0.0394	0.0657	0.1677
4	0.0217			
$T_1 = 100$				
0	0.1560	0.1851	0.2494	0.3229
1	0.1106	0.1275	0.1442	0.2035
2	0.0775	0.0810	0.1132	0.1038
3	0.0354	0.0403		
4	0.0193			

$$\eta_\tau^0(t) \approx 1 - \frac{\frac{u_{3\tau}^{k+1} - u_{3\tau}^k}{\tau_k} + \delta u_{3\tau}^k}{\alpha u_{2\tau}^k}. \tag{97}$$

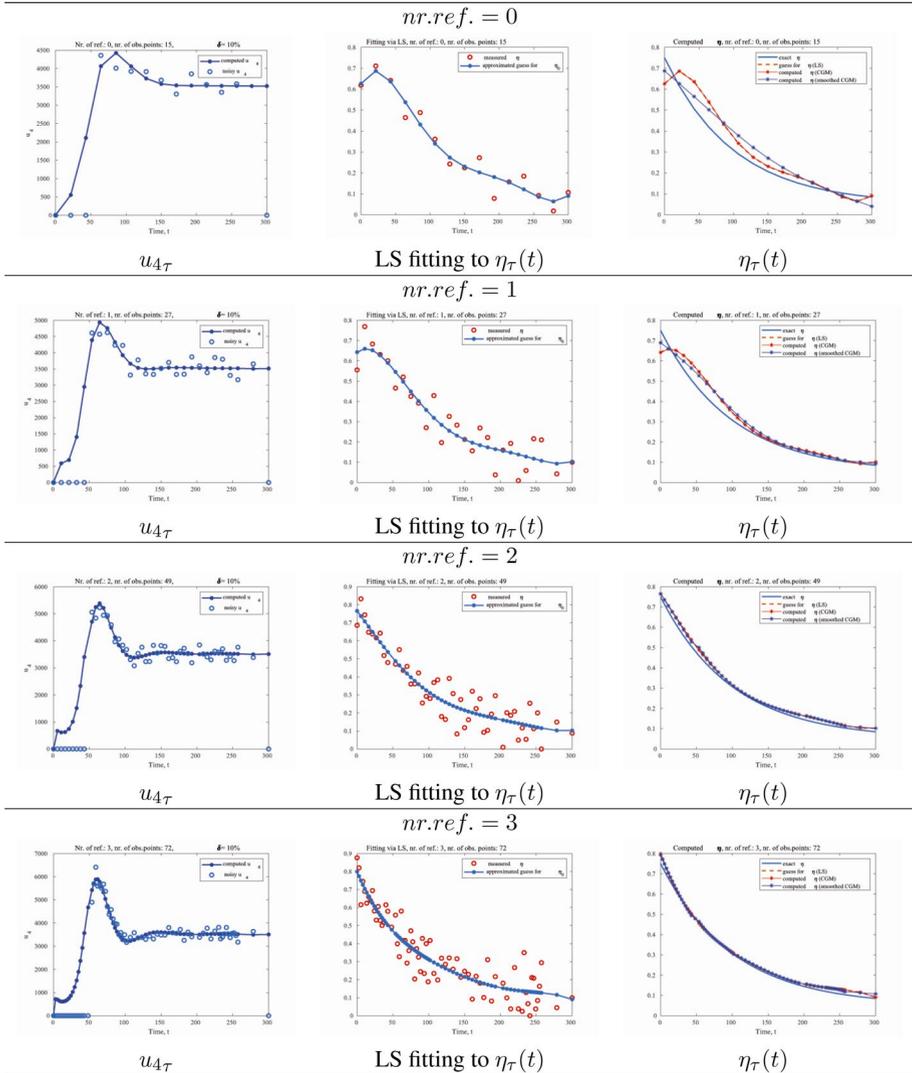
Here,  $u_{3\tau}^{k+1}, u_{3\tau}^k, u_{2\tau}^k$  are known computed approximations of functions  $u_3, u_2$  at time iterations  $k + 1$  and  $k$ , respectively. In our computations we take values of  $u_{3\tau}^{k+1}, u_{3\tau}^k, u_{2\tau}^k$  using solution of the problem (1)–(2) with exact values of the test model function  $\eta(t)$  and then adding the noise  $\delta$  as

$$u_{i\sigma}(t) = u_{i\sigma}(t)(1 + \sigma\alpha), \quad i = 2, 3, \tag{98}$$

where  $\sigma \in [0, 1]$  is nose level and  $\alpha \in [-1, 1]$  is random number. We note that denominator in (97) is not approaching zero because  $\alpha = 0.4$  and  $u_{2\tau}(t) > 0 \forall t \in [0, T]$ . To get reasonable approximation  $\eta_\tau^0$  for the initial guess  $\eta^0$  in Algorithm 2 we assume that noisy functions  $u_{3\sigma}, u_{2\sigma}$  are known on the initial non-refined mesh, apply (97) and then use polynomial fitting to obtained noisy data  $\eta_\tau^0$  in order smooth them. Finally, the condition (5) was applied for the computed  $\eta_\tau^0$  in order to ensure that  $\eta^0(t)$  belongs to the set of admissible parameters  $M_\eta$ . Second order discretization of the first time derivative in (97) is also possible. We note that numerical differentiation of noisy data is an ill-posed problem and it is discussed in detail in Section 4 of [25].

All tests are performed with tolerance  $\Theta = 10^{-7}$  in ACGA algorithm and  $\beta_1 = 0.1$  in (8.94). The value of  $\beta_1$  is chosen such that it allows local refinements and avoids refinement of the very large time region in the time mesh. All tests are performed for different  $T_1 = 25, 50, 100$  for the time interval  $[T_1, T_2] = [T_1, 300]$  which corresponds to the fact that HIV virus can be detected in the first 3-9 weeks after infection.

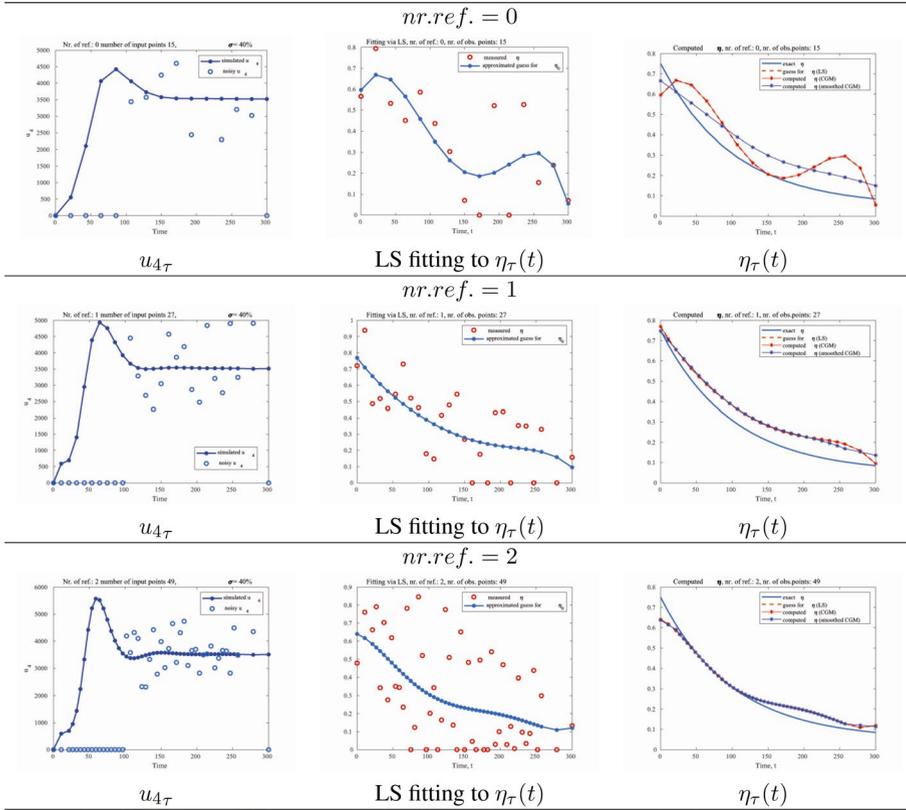
Relative errors in the reconstructed parameters  $\eta(t)$  presented in the Tables are measured in  $L_2$ -norm and are computed as



**Fig. 4** Test 1. Left figures: simulated  $u_{4\tau}$  vs. noisy  $u_{4\tau\sigma}$  on the different adaptively refined time meshes. Here, noisy observed data are presented by circles. Middle figures: least squares fitting for the noisy  $\eta_\tau$ . Right figures: results of ACGA on adaptively refined meshes. Computations are done for noise level  $\sigma = 10\%$  in  $u_4$  and for  $T_1 = 50$

$$e_\eta = \frac{\|\eta - \eta_\tau\|_{L_2(\Omega_T)}}{\|\eta\|_{L_2(\Omega_T)}}. \tag{99}$$

Complete description of all numerical tests with reconstruction results are presented in the recent work [11].



**Fig. 5** Test 1. Left figures: simulated  $u_{4\tau}$  vs. noisy  $u_{4\tau\sigma}$  on the different adaptively refined time meshes. Here, noisy observed data are presented by circles. Middle figures: least squares fitting for the noisy data  $\eta_\tau$ . Right figures: results of ACGA on adaptively refined meshes. Computations are done for noise level  $\sigma = 40\%$  in  $u_4$  and for  $T_1 = 100$

**Test 1**

See Table 4.

In this test we present the reconstruction results of the smooth model function

$$\eta(t) = 0.7e^{-t} + 0.05, t \in [0, 300] \tag{100}$$

for different starting time points  $T_1 = 25, 50, 100$  and for the number of discretization points  $k = 15$  on the initial time partition  $\mathcal{J}_\tau^0$  which is generated with equidistant time step  $\tau = 300/(k - 1)$ . More precisely, in this test we model the control parameter  $\eta(t)$  as a smooth function given by the equation (100), and we want to recover this function on the whole time interval  $[0, T]$  using measurements of the noisy virus population function  $u_{4\tau\sigma}$  for different values of the initial measurement of this function, or for different times  $T_1$ .

Left figures of Figs. 4 and 5 show simulated  $u_{4\tau}$  versus noisy  $u_{4\tau\sigma}$  on the different adaptively refined in time meshes. On these figures noisy observed data is presented by blue

**Table 5** Test 2. Relative errors  $e_\eta$  computed for reconstruction of the function  $\eta(t) = 0.7, t \in [0, 300]$  for  $T_1 = 25, 50, 100$  on different locally adaptively refined time-meshes

$\sigma$ nr.of ref.	5 %	10%	20%	40%
$T_1 = 25$				
0	0.0718	0.0802	0.0834	0.0617
1	0.0592	0.0315	0.0290	0.0493
2	0.0403	0.0091		0.0301
3	0.0272	0.0050		0.0240
4	0.0191	0.0064		
5	0.0170			
6	0.0117			
$T_1 = 50$				
0	0.0725	0.0758	0.0720	0.1026
1	0.0656	0.0572	0.0694	0.0730
2	0.0459	0.0414	0.0505	0.0571
3	0.0273	0.0239	0.0179	0.0236
4	0.0111	0.0183		
5	0.0066	0.0099		
$T_1 = 100$				
0	0.0801	0.0676	0.0535	0.0852
1	0.0568	0.0547		0.0487
2	0.0351	0.0481		0.0208
3	0.0265	0.0265		
4	0.0212	0.0130		
5	0.0095	0.0090		
6	0.0084			

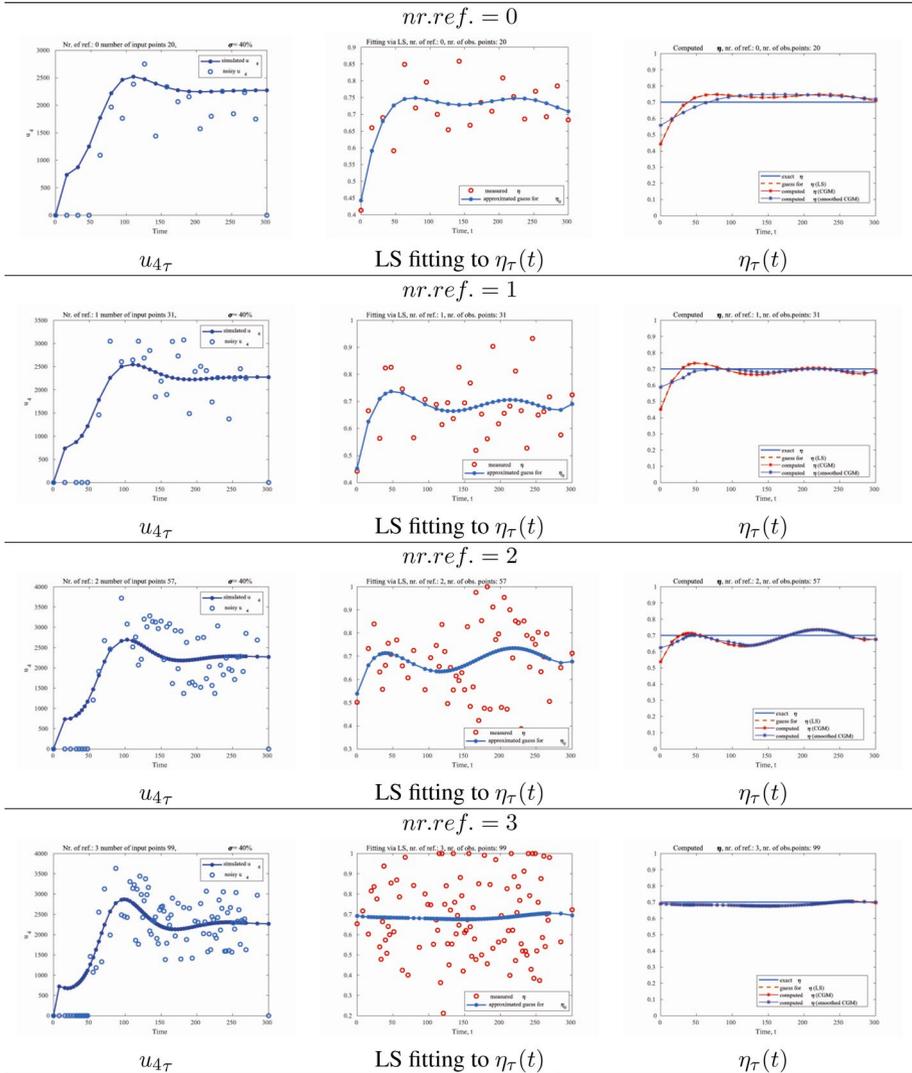
circles and simulated data without noise is shown by the solid blue line. Middle figures of Figs. 4 and 5 present least squares fitting for the noisy  $\eta_\tau$ . We recall that initial guess for noisy  $\eta_\tau$  is computed via applying (97) and is represented by the red circles on the middle figures of Figs. 4 and 5. The least squares fitting to the noisy  $\eta_\tau$  is shown by the solid blue line on these figures.

Results of reconstruction of the model function (100) for noise levels  $\sigma = 5\%, 10\%, 20\%, 40\%$  in virus population function  $u_4(t)$  are presented Table 1. Right figures of Fig. 4 show the reconstruction results of the function (100) for a noise level  $\sigma = 10\%$  in the data  $u_4(t)$  for starting observed time  $T_1 = 50$ . Right figures of Fig. 5 shows the reconstruction results of this function for a noise level  $\sigma = 40\%$  in the data  $u_4(t)$  and for  $T_1 = 100$ .

Table 1 and Figs. 4 and 5 confirm that with local time-mesh refinements the reconstruction of the drug efficacy function  $\eta_\tau$  is significantly improved compared to the reconstruction of  $\eta_\tau$  obtained on the initial non-refined time-mesh.

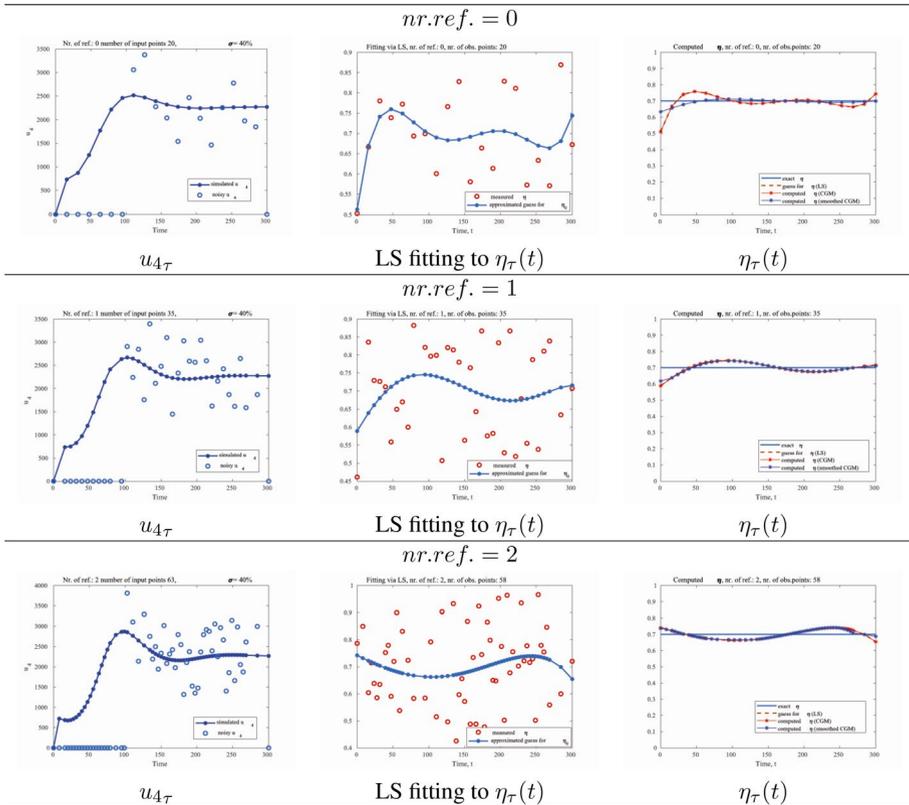
## Test 2

See Table 5.



**Fig. 6** Test 2. Left figures: simulated  $u_{4T}$  vs. noisy  $u_{4T\sigma}$  on the different adaptively refined time meshes. Here, noisy observed data are presented by circles. Middle figures: least squares fitting for the noisy  $\eta_T$ . Right figures: results of ACGA on adaptively refined meshes. Computations are done for noise level  $\sigma = 40\%$  in  $u_4$  and for  $T_1 = 50$

In this test we present numerical reconstruction results of the constant model function  $\eta(t) = 0.7$  from noisy observations of the virus population function  $u_4(t)_\sigma$  at the observation interval  $[T_1, T_2]$ . We again took  $T_1 = 25, 50, 100$ , but number of observation points were 20 at the time interval  $[T_1, T_2] = [T_1, 300]$ . We generate initial time partition  $\mathcal{J}_T$  with equidistant time step  $\tau = 300/19$ . The reconstruction results of the model function  $\eta(t) = 0.7$  for noise levels  $\sigma = 5\%, 10\%, 20\%, 40\%$  in data  $u_4(t)$  are presented in Table 2. Right figures of



**Fig. 7** Test 2. Left figures: simulated  $u_{4\tau}$  vs. noisy  $u_{4\tau\sigma}$  on different adaptively refined time meshes. Here, noisy observed data are presented by circles. Middle figures: least squares fitting to noisy data for  $\eta_\tau$ . Right figures: results of ACMA on adaptively refined meshes. Computations are done for noise level  $\sigma = 40\%$  in  $u_4$  and for  $T_1 = 100$

Figs. 6, 7 show the reconstruction results of the function  $\eta(t) = 0.7$  for noise level  $\sigma = 40\%$  in the data  $u_4(t)$  for  $T_1 = 50$  and  $T_1 = 100$ , respectively.

We again observe from the results of Table 2 and Figs. 6 and 7 that with local time-mesh refinements the reconstruction of the drug efficacy  $\eta_\tau$  is significantly improved compared to the reconstruction of  $\eta_\tau$  obtained on the initial non-refined time-mesh even if we add large noise  $\sigma = 40\%$  to the observed data  $u_4(t)$ .

### Conclusion

The finite element time-adaptive optimization method for determination of the drug efficacy in a mathematical model of HIV infection with drug therapy is presented. Time-adaptive optimization means that first the time-dependent drug efficacy is determined at a known coarse time partition using several known values of observed function (usually, we used 15–20 observations). Then the time-mesh is locally refined at points where the computed residual  $|R(\eta_\tau)|$  attains its maximal values and the drug efficacy is computed

on a new refined time-mesh until the relative error in the reconstructed parameter  $\eta$  is reduced to the desired accuracy. Numerical experiments show efficiency and reliability of proposed adaptive method on reconstruction of different model functions  $\eta$  from noisy observed virus population function.

The proposed new time-adaptive method can eventually be used by clinicians to determine the drug-response for each treated individual. The exact knowledge of the personal drug efficacy can aid in the determination of the most suitable drug as well as the most optimal dose for each person, in the long run resulting in a personalized treatment with maximum efficacy and minimum adverse drug reactions.

The proposed time-adaptive method can be adopted to solve multi-parameter identification problems for a broad class of problems stated by the system of ODE.

**Acknowledgements** The research of L. B. is supported by the Swedish Research Council grant VR 2018-03661. The research of I. G. author was supported by the Russian Foundation for Basic Research (grant 20-01-00352).

**Funding** Open access funding provided by University of Gothenburg.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Albersmeyer, J.: Adjoint based algorithms and numerical methods for sensitivity generation and optimization of large scale dynamic systems, Ph.D. thesis, Heidelberg University (2010)
2. Almudevar, A., Arruda, E.F.: Optimal approximation schedules for a class of iterative algorithms, with an application to multigrid value iteration. *IEEE Trans. Autom. Control* **57**(12) (2012)
3. Arruda, E.F., Dias, C.M., de Magalhaes, C.V., Pastore, D.H., Thomé, R.C.A., Yang, H.M.: An optimal control approach to HIV immunology. *Appl. Math.*, 1115–1130 (2015). <https://doi.org/10.4236/am.2015.66102>
4. Bakushinskii, A.B., Kokurin, MYu.: *Iterative Methods for Approximate Solution of Inverse Problems*. Springer, New York (2004)
5. Bangerth, W., Joshi, A.: Adaptive finite element methods for the solution of inverse problems in optical tomography. *Inverse Probl.* **24**, 034011 (2008)
6. Banks, H.T., Kunisch, K.: *Estimation techniques for distributed parameter systems*. Birkhäuser (1989). ISBN 978-1-4612-3700-6
7. Becker, R., Rannacher, R.: An optimal control approach to a posteriori error estimation in finite element method. *Acta Numer.* **10**, 1–102 (2001)
8. Beilina, L.: Adaptive finite element method for a coefficient inverse problem for the Maxwell's system. *Appl. Anal.* **90**, 1461–1479 (2011)
9. Beilina, L., Gainova, I.: Time-adaptive FEM for distributed parameter identification in biological models. *Appl. Inverse Probl. Springer Proc. Math. Stat.* **48**, 37–50 (2013)
10. Beilina, L., Gainova, I.: Time-adaptive FEM for distributed parameter identification in mathematical model of HIV infection with drug therapy. *Inverse Probl. Appl. Springer Proc. Math. Stat.* **120**, 111–124 (2015)
11. Beilina, L., Gainova, I.: Time-adaptive optimization in a parameter identification problem of HIV infection. [arXiv:1912.01112](https://arxiv.org/abs/1912.01112)
12. Beilina, L., Johnson, C.: A posteriori error estimation in computational inverse scattering. *Math. Models Methods Appl. Sci.* **15**, 23–37 (2005)

13. Beilina, L., Klibanov, M.V.: *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*. Springer, New York (2012)
14. Beilina, L., Klibanov, M.V., Kokurin, MYu.: Adaptivity with relaxation for ill-posed problems and global convergence for a coefficient inverse problem. *J. Math. Sci.* **167**, 279–325 (2010)
15. Bocharov, G., Chereshev, V., Gainova, I., Bazhan, S., Bachmetyev, B., Argilagué, J., Martínez, J., Meyerhans, A.: Human immunodeficiency virus infection: from biological observations to mechanistic mathematical modelling. *Math. Model. Nat. Phenom.* **7**(5), 78–104 (2012)
16. Bocharov, G., Volpert, V., Ludewig, B., Meyerhans, A.: *Mathematical Immunology of Virus Infections*. Springer/Nature (2018). ISBN 978-3-319-72317-4
17. Burden, R.L., Douglas Fairés, J.: *Numerical Analysis*, 9th edn. Brooks/Cole
18. Chereshev, V.A., Bocharov, G.A., Bazhan, S.I., Bachmetyev, B., Gainova, I.A., Likhoshvai, V.A., Argilagué, J.M., Martínez, J.P., Rump, J.A., Mothe, B., Brander, C., Meyerhans, A.: Pathogenesis and treatment of HIV infection: the cellular, the immune system and the neuroendocrine systems perspective. *Int. Rev. Immunol.* **32**(3), 282–306 (2013). <https://doi.org/10.3109/08830185.2013.779375>
19. Chereshev, V.A., Bocharov, G.A., Kim, A.V., Bazhan, S.I., Gainova, I.A., Krasovskii, A.N., Shmagel, N.G., Ivanov, A.V., Safronov, M.A., Tretyakova, R.M.: *Introduction to Modeling and Control of HIV Infection Dynamics*. Institute for Computer Research, Izhevsk (2016) (**book, in Russian**)
20. Deutsch, F.: Existence of best approximations. *J. Approx. Theory* **28**, 132–154 (1980)
21. VI Eastern Europe and Central Asia AIDS Conference. <http://en.eecaac2018.org/>
22. Eftimie, R., Gillard, J.J., Cantrell, D.: Mathematical models for immunology: current state of the art and future research directions. *Bull. Math. Biol.* **78**(10), 2091–2134 (2016)
23. Eriksson, K., Estep, D., Johnson, C.: *Calculus in Several Dimensions*. Springer, Berlin (2004)
24. Eriksson, K., Estep, D., Hansbo, P., Johnson, C.: *Computational Differential Equations*. Cambridge University Press, Cambridge (1996)
25. Eriksson, M.: *Parameter identification in a mathematical model of HIV infection with drug therapy*, Master's thesis. <http://hdl.handle.net/2077/54664>
26. Hatz, K., Schlöder, J.P., Bock, H.G.: Estimating parameters in optimal control problems. *SIAM J. Sci. Comput.* **34**(3) (2012). <https://doi.org/10.1137/110823390>
27. Kepler, T.B., Perelson, A.S.: Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11514–11519 (1998)
28. Klibanov, M.V., Bakushinsky, A.B., Beilina, L.: Why a minimizer of the Tikhonov functional is closer to the exact solution than the first guess. *J. Inverse Ill-Posed Probl.* **19**, 83–105 (2011)
29. Koshev, N., Beilina, L.: An adaptive finite element method for Fredholm integral equations of the first kind and its verification on experimental data. *CEJM* **11**(8), 1489–1509 (2013)
30. Manky, L.M., Temin, H.M.: Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* **69**, 5087–5094 (1995)
31. Matlab source code AFEM\_HIV. [https://github.com/larisa70/AFEM\\_HIV](https://github.com/larisa70/AFEM_HIV)
32. Nowak, M.A., Bangham, C.R.M.: Population dynamics of immune responses to persistent viruses. *Science* **272**, 74–79 (1996). <https://doi.org/10.1126/science.272.5258.74>
33. Nowak, M.A., May, R.: *Virus Dynamics Mathematical Principles of Immunology and Virology*. Oxford University Press, Oxford (2000)
34. Nowak, M.A., Bonhoeffer, S., Shaw, G.M., May, R.M.: Anti-viral drug treatment: dynamics of resistance in free virus and infected cell populations. *J. Theor. Biol.* **184**, 203–217 (1997)
35. Ouifki, R., Witten, G.: A model of HIV-1 infection with HAART therapy and intracellular delay. *Discrete Contin. Dyn. B* **8**, 229–240 (2007)
36. Pastore, D.H., Thomé, R.C.A., Dias, C.M., Arruda, E.F., Yang, H.M.: A model for interactions between immune cells and HIV considering drug treatments. *Comput. Appl. Math.* **37**, 282–295 (2018). <https://doi.org/10.1007/s40314-017-0528-8>
37. Patrick, G.L.: *An Introduction to Medicinal Chemistry*. Oxford University Press, Oxford (2013)
38. Polak, E.: *Optimisation. Algorithms and Consistent Approximations*, vol. 124. AMS, Springer
39. Rong, L., Gilchrist, M.A., Feng, Z., Perelson, A.S.: Modeling within host HIV-1 dynamics and the evolution of drug resistance: trade-offs between viral enzyme function and drug susceptibility. *J. Theor. Biol.* **247**, 804–818 (2007)
40. Shu, H., Wang, L., Watmough, J.: Sustained and transient oscillations and chaos induced by delayed anti-viral immune response in an immunosuppressive infection model. *J. Math. Biol.* **68**, 477–503 (2014)
41. Smith, R.J., Wahl, L.M.: Distinct effects of protease and reverse transcriptase inhibition in an immunological model of HIV-1 infection with impulsive drug effects. *Bull. Math. Biol.* **66**, 1259–1283 (2004)
42. Srivastava, P.K., Banerjee, M., Chandra, P.: Modeling the drug therapy for HIV infection. *J. Biol. Syst.* **17**(2), 213–223 (2009)

43. Teschl, G.: *Ordinary Differential Equations and Dynamical Systems*, Graduate Studies in Mathematics, vol. 140. American Mathematical Society, Providence (2012)
44. Tikhonov, A.N.: On the stability of inverse problems (in Russian). *Doklady USSR Acad. Sci.* **39**, 195–198 (1943)
45. Tikhonov, A.N., Arsenin, V.Ya.: *Solutions of Ill-Posed Problems*. Winston and Sons, Washington (1977)
46. Tikhonov, A.N., Goncharsky, A.V., Stepanov, V.V., Yagola, A.G.: *Numerical Methods for the Solution of Ill-Posed Problems*. Kluwer, London (1995)
47. UNAIDS. <http://aidsinfo.unaids.org>
48. Yan, A.W.C., Cao, P., McCaw, J.M.: On the extinction probability in models of within-host infection: the role of latency and immunity. *J. Math. Biol.* **73**, 787–813 (2016)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.