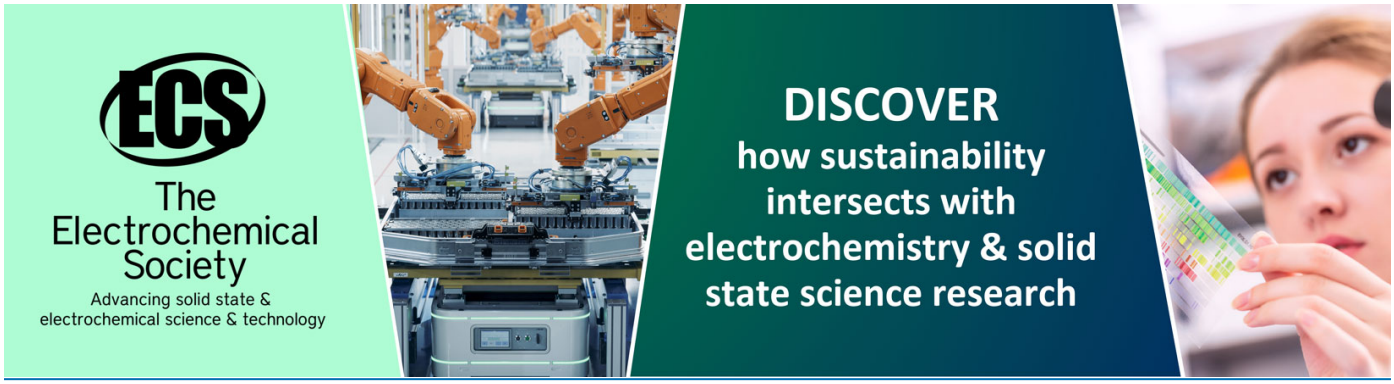
Smooth-polylines vs. polynomials in balanced identification technology: solving inverse problems of thermal conductivity



Smooth-polylines vs. polynomials in balanced identification technology: solving inverse problems of thermal conductivity

A V Sokolov* **and V V Voloshinov**

Institute for information transmission problem (Kharkevitch Insitute) RAS, Bolshoy Karetny per. 19, build.1, Moscow 127051 Russia

*E-mail: alexander.v.sokolov@gmail.com

Abstract. Various parameters' identification problems of one-dimensional nonlinear heat equation are considered. Their numerical study was carried out on the basis of balanced identification technology, which provides a compromise between the simplicity of the model (the curvature of the functions) and the proximity to experimental data. The problem of identifying functions whose arguments are model variables is considered. When approximating such a function, we had to use a polynomial function – the use of polylines (polygonal lines) in this case (superposition of functions) leads to nonsmooth mathematical programming problems (with discontinuous derivatives) with the solution not supported by standard solvers. An investigation of the use of special smooth approximations of polygonal curves (smooth-polylines) is presented.

1. Introduction

The experiment preparation and processing of the results involve an extensive use of mathematical models of the objects under study. To save costs they must be carefully planned: one should determine what, when, where and with what accuracy is to be measured to estimate the sought parameters with given accuracy. These questions can be answered by "rehearsing" the experiment and its processing on a mathematical model that simulates the behaviour of the object.

Usually an experiment's purpose is to evaluate some of the object's parameters. In case of an indirect experiment some parameters are measured, while others are to be evaluated. The relationship between the parameters can be described by complex mathematical models. The formalization of this approach leads to identification problems that are inverse by nature. Those problems often turn out to be ill-posed and specific approaches using regularization methods are required for the solution [1]. In the balanced identification (SvF) method [2] the choice of regularization weights is making by minimizing the cross-validation error. It enables to find a balanced solution giving the optimal (in the sense of minimizing the cross-validation error) compromise between the proximity of the model to the data and the simplicity of the model [3], formalized in criterion by regularizing additive.

SvF-technology has been successfully used by solving inverse problems in various scientific fields (mechanics, plasma physics, biology, plant physiology, epidemiology, meteorology, atmospheric pollution transfer, etc., a more detailed enumeration can be found in [2]).

Another problem is the smoothness of the criterion. The use of NLP-solvers, as a rule, leads to the requirement of continuous differentiability of the solution selection criterion, which in turn entails continuous differentiability of functions, including those that depend on other model variables.

Polylines are widely used in mathematical modeling. They, unlike polynomials or Fourier series, have the property of locality - the value of an individual parameter (node) affects the function being approximated only up to neighboring nodes. However, a significant disadvantage of polylines is the discontinuities of derivatives at the grid nodes.

The study of methods for smoothing polylines and the possibility of their use in solving inverse problems of processing experiments is the purpose of this work. The efficiency of the developed approximation methods is demonstrated by solving inverse problems of thermal conductivity.

The main task is to find the dependence of the thermal conductivity coefficient on temperature based on an array of experimental data. Additionally, knowledge about the object will be involved: the equation of thermal conductivity.

2. Smooth-polylines in SvF-method

One of the SvF-method issues arises if the model contains composition of unknown functions. The typical case is differential equations. Let's take autonomous differential equation of the 1st order :

$$x'(t) = F\left(x(t)\right) \tag{1}$$

Assume that we have dataset $\{t_i, x_i\}_{i=0}^{I}$, (points of solution of (1), may be, with errors) and want to identify both functions $x(t), F(x)$ by SvF-method with respect to the following "domain" constraints

$$x(\cdot):[0,T]\to[X_{LO},X_{UP}],\ \ F(\cdot):[X_{LO},X_{UP}]\to[F_{LO},F_{UP}]$$

In the above simple case SvF-method iteratively solves some finite dimensional analogue, actually – mathematical programming problem (MPP), of the following variational problem (α is a regularization coefficient to be tuned by SvF-method further)

$$\begin{aligned}&\sum_{i=0:I}\left(x_i - x(t_i)\right)^2 + \alpha\int_{X_{\min}}^{X_{\max}}\left(F''(x)\right)^2 dx \xrightarrow[(x(\cdot),F(\cdot))]{} \min,\ \text{w.r.t.}\\ &x'(t) = F\left(x(t)\right),\ x(\cdot):[0,T]\to[X_0,X_1],\ \ F(\cdot):[X_0,X_1]\to[F_0,F_1].\end{aligned} \tag{2}$$

There are different ways to get that MPP, for sake of brevity we consider one of that. The first step of discretization is to represent function $x(t)$ as a set of unknown values $x_k = x(t_k),\ k = 0:N_T$ on a fixed mesh $\{t_k\}\subseteq[T_{LO},T_{UP}]$. Take uniform mesh $t_k = T_{LO} + k\left(T_{UP}-T_{LO}\right)N_T^{-1},\ k = 0:N_T$, and replace continues function *x(t)* with piecewise linear function (*pw-function*) $\tilde{\ }$ $\}_{k=0}^{N_T}\Big)$ passing through points $\left\{\left(t_k,x_k\right)\right\}_{k=0}^{N_T}$. As to discretization of *F(x)* we consider two approaches. The first is to approximate function $F(x)$ by polynomial with unknown coefficients, i.e. $\tilde{I}$ $\overline{p=0:P}\ c_p x^p$. Thus variational problem (2) may be replaced with MPP:

$$\begin{aligned}&\sum_{i=0:I}\Big(x_i - \tilde{\ }\ \Big)\Big)^2 + \alpha\int_{X_{\min}}^{X_{\max}}\left(\sum_{p=2:P}c_p p\left(p-1\right)x^p\right)^2 dx \xrightarrow[\left(\{x_k\}_{k=0}^{N_T},\ \{c_p\}_{p=0}^{P}\right)]{} \min,\ \text{where}\\ &\frac{x_k - x_{k-1}}{\Delta t} = \sum_{p=0:P}c_p\left(\frac{x_k + x_{k-1}}{2}\right)^p,\ t_k = T_{LO} + k\Delta t,\ \Delta t = \frac{T_{UP}-T_{LO}}{N_T}, k = 0:N_T,\ \ \left\{c_p\right\}_{p=0}^{P}\in C\subset\mathbb{R}\end{aligned} \tag{3}$$

Polylines or (smoothed) piecewise approach is based on another approximation of $F(x)$ - by piecewise continuous function defined on the interval $[X_{LO},X_{UP}]$ by a finite set $\left\{\xi_j,F_j\right\}_{j=0}^{N_X}\subseteq[X_{LO},X_{UP}]\times[F_{LO},F_{UP}]$ of unknown values $F_j = F(\xi_j)$ on a fixed mesh over x variable.

Let $A_j = \left(F_j - F_{j-1}\right)\left(\xi_j - \xi_{j-1}\right)^{-1}, j = 1: N_X$, be notation for the slopes of pw-function on intervals of its domain. Now pw-approximation of *F(x)* may be defined as follows (hereinafter $[z]_+ = \max\{z, 0\}$)

$$\tilde{\imath} \qquad + A_1\left(x - \xi_0\right) + \sum_{j=2}^{N_x}\left(A_j - A_{j-1}\right)\left[x - \xi_{j-1}\right]_+ \tag{4}$$

The disadvantageous of the expression (4) is that the function $\tilde{\imath}$ is not differentiable and appropriate MPP, "close" to (2), could not be solved by available NLP-solvers (e.g. IPOPT, https://coin-or.github.io/Ipopt). The issue of smoothing functions like $[z]_+$ has a long history [4, 5]. The most popular are the following three approximations, all dependent on small parameter ε:

$$[z]_+ \approx 0.5\left(z + \sqrt{z^2 + \varepsilon}\right),\ [z]_+ \approx 0.5\left(z + \frac{z^2}{\sqrt{z^2 + \varepsilon}}\right),\ [z]_+ \approx z + \varepsilon \ln\left(\exp\left(\frac{-z}{\varepsilon}\right) + 1\right)$$

In the current implementation of SvF-technology we used the 1st, so called CHKS (Chen-Harker-Kanzow-Smale) function, because (in some sense) it gives non-linear mathematical programming problem with (only) 2nd order polynomial in equation constraints. The 2nd option is quite similar and even more accurate because equal to 0 when $z=0$. After that expression (4) is replaced with the follow

$$\begin{gathered}\Phi(x, \varepsilon) = F_0 + A_1\left(x - \xi_0\right) + \frac{1}{2}\sum_{j=2}^{N_x}\left(A_j - A_{j-1}\right)\left(x - \xi_{j-1} + \eta_{j-1}(x, \varepsilon)\right), \\ \eta_j(x, \varepsilon) = \sqrt{\left(x - \xi_j\right)^2 + \varepsilon}, \text{ or } \eta_j(x, \varepsilon) = \left(x - \xi_j\right)^2\left(\left(x - \xi_j\right)^2 + \varepsilon\right)^{-\frac{1}{2}}.\end{gathered} \tag{5}$$

In the case of uniform mesh by x variable, $x_j = X_0 + j\Delta x, j = 0: N_X, \Delta x = \left(X_1 - X_0\right)N_X^{-1}$, expression (5) may be reduced to the following

$$\Phi(x, \varepsilon) = \frac{1}{2}\left(F_0 + F_{N_X} + \frac{F_1 - F_0}{\Delta x}\left(x - \xi_0\right) + \frac{F_{N_X} - F_{N_X-1}}{\Delta x}\left(x - \xi_{N_X}\right) + \sum_{j=2}^{N_x}\left(\frac{F_j - 2F_{j-1} + F_{j-2}}{\Delta x}\right)\eta_{j-1}(x, \varepsilon)\right)$$

Now the optimization problem (2) (corresponding to the "demo" inverse problem for autonomous 1st order ODE) will be the following (for some regularization coefficient α and the fixed parameter ε):

$$\begin{gathered}\sum_{i=0:I}\left(x_i - \tilde{\ } \qquad \}_{k=1}^{N_T})\right)^2 + \alpha \sum_{j=1:N_X-1}\frac{\left(\Phi\left(\xi_{j+1}, \varepsilon\right) - 2\Phi\left(\xi_j, \varepsilon\right) + \Phi\left(\xi_{j-1}, \varepsilon\right)\right)^2}{\Delta x^3} \xrightarrow[\left(\{x_k\}_{k=1}^{NT}, \{F_j\}_{j=0}^{NX}\right)]{} \min, \\ t_k = T_{LO} + k\Delta t,\ \Delta t = \frac{T_{UP} - T_{LO}}{N_T}, k = 0: N_T, \quad \xi_j = X_{LO} + j\Delta x,\ \Delta x = \frac{X_{UP} - X_{LO}}{N_X}, j = 0: N_X, \\ \frac{x_k - x_{k-1}}{\Delta t} = \Phi\left(\frac{x_k + x_{k-1}}{2}, \varepsilon\right),\ X_{LO} \le x_k \le X_{UP},\ k = 1: N_T,\ F_{LO} \le F_j \le F_{UP},\ j = 0: N_X.\end{gathered}$$

3. One-dimensional thermal conductivity problem with initially and boundary conditions

Let's denote "M=0" a set of mathematical statements defining the model of thermal conductivity:

$$M = 0\colon \begin{cases} x \in [0,2], t \in [0,5] \\ \frac{\partial T}{\partial t} = \frac{\partial}{\partial x}\left(K(T)\frac{\partial T}{\partial x}\right) \\ T(x, 0) = \varphi(x) \\ T(0, t) = l(t) \\ T(2, t) = r(t) \end{cases} \tag{6}$$

where x and t are the spatial and temporal coordinates, T(x,t) is the temperature (°C), K(T) is the thermal conductivity coefficient (temperature-dependent), φ(t) is the initial condition, l(t) and r(t) are the left and right boundary conditions.

When conducting numerical experiments, an exact solution is used to generate datasets (of observations) and to calculate the errors of the solution:

$$Ts(x,t) = \frac{200(t+1)}{(x+1)^2 + (t+1)^2}; \quad Ks(T) = \frac{100}{T}$$

The functions of the exact solution are shown in figure 1.

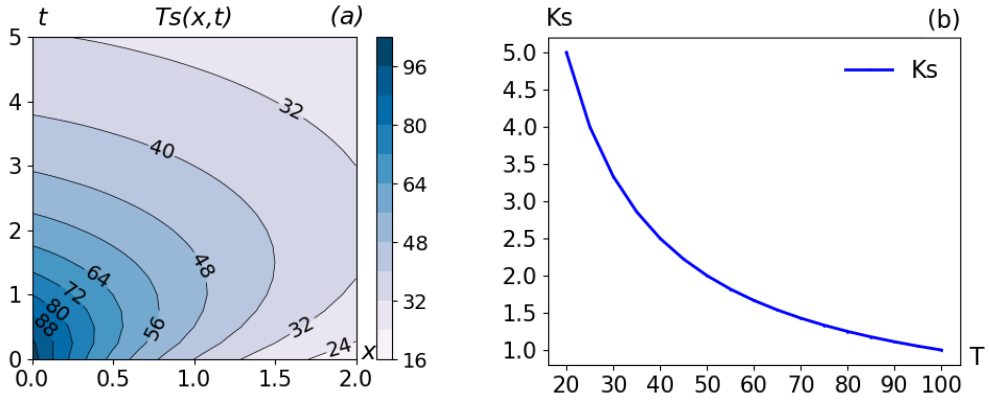


Figure 1. Functions of the exact solution: (a) contour lines of *Ts(x,t)*; (b) thermal conductivity *K(T)*.

4. Data set

We formalize the concept of a data set (observations or measurements set):

$$D: \{x_i, t_i, T_i, \}, i \in I, \; I = 0..i_{max},$$

where Ti is the temperature measurement at point xi at time ti.

For vectors of dimension |*D*|, we introduce the notation

$$\|a_i\|_D = \|a\|_D = \left(\frac{1}{|D|}\sum_{i \in I} a_i^2\right)^{1/2}$$

Below, for numerical experiments, a data (measurements) set is used, generated on a regular 11x11 grid (11 points in space 0, 0.2, 0.4 ..., 2 and 11 points in time 0, 0.5, 1, ... 5)

$$D: \; \{x_i = n * 0.2, \; t_i = j * 0.5, \; T_i = Ts(x_i, t_i) + \varepsilon_i\}, i = 11 * j + n, \qquad n = 0..10, \qquad j = 0..10,$$

where $Ts(x_i,t_i)$ are the values of the exact solution, ε_i is the random error with variance

$$\sigma_d = \|\varepsilon\|_D.$$

To generate ε_i, a normal distribution random number generator (gauss (0.2)) with zero mean and variance equal to 2 (°C) was used. As a result, the distribution ε_i was obtained with average md = -0.10 (°C) and variance σ_d = 2.06 (°C). These characteristics of errors are not used in calculations but are taken into account when considering the results. The location of the measurement points of the D set on the x, t plane can be seen in figures 2, 3, 4. The data set files can be found in file SvF-2021-11.zip in the Git repository https://github.com/distcomp/SvF.

5. Method of balanced identification (SvF)

The balanced identification method is described in details in [2].

We formalize the concepts necessary for its use in this problem:

- a measure of the proximity of the model trajectory to measurements (data set D) or the approximation error: $\mathrm{MSD(D,T)} = \frac{1}{|D|}\sum_{i\in I}\left(T_i - T(x_i,t_i)\right)^2 = \|T_i - T(x_i,t_i)\|_D^2$,

where $|D|$ – the number of elements of the set D;

- a measure of curvature (complexity) of functions of one variable $f(x)$

$$\mathrm{Curv}(f(x),\alpha) = \alpha\int_a^b (f''(x))^2 dx\,,$$

where $[a, b]$ is the domain of the function;

- a objective function - a combination of the measures introduced above

$$\mathrm{Obj} = \mathrm{MSD(D,T)} + \mathrm{Curv}(K(T),\alpha_T) + \mathrm{Curv}\left(\varphi(x),\alpha_\varphi\right) + \mathrm{Curv}(l(t),\alpha_l) + \mathrm{Curv}(r(t),\alpha_r),$$

where $K(T)$, $\varphi(x)$, $l(t)$ и $r(t)$ – functions defining solution (6).

Curvature terms (last four) are the regularizing additions that makes the problem (of the search for a continuous functions) correct. The choice of its values determines the quality of the solution. The choice is carried out by minimizing the root-mean-square error of cross-validation [2].

6. Various identification problems and their numerical solution

To find approximate solutions, we use numerical models, that are obtained from analytical ones by replacing of continues functions with either grid-functions or polynomials/smooth-polylines (only for $K(T)$), derivatives - with finite differences, integrals - with the sums. Note that the grid used for numerical model (41 points in x with a step equal to 0.05 and 21 points in t with a step equal to 0.25) is not tied to the measurement points in any way. For simplicity (and stability of calculations), an implicit four-point scheme [6] was used for approximation of temperature derivatives.

The following notations are used for errors below:

σ^*_{cv} – error of cross-validation (°C), the main indicator of the "quality" of the constructed model;

rmsd* - standard deviation (°C) of the SvF solution from observations,

Δ - standard deviation (°C) of the SvF solution from the exact one.

6.1. K(T) - 7th degree polynomial and least square fitting

Let's solve the problem of finding unknown functions by the least squares method (without regularization). See figure 2.

Errors (°C): σ^*_{cv}=3.52, rmsd*= 1.05, Δ=2.38.

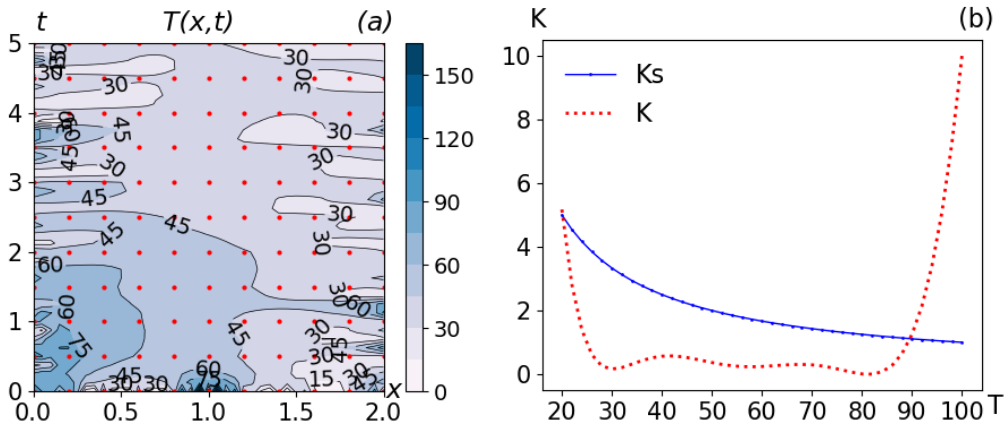


Figure 2. MSD-minimization solution of Problem 6.1. (a) contour lines of the solution $T(x,t)$; (b) thermal conductivity: K – solution, Ks – exact solution.

6.2. K(T) - 7th degree polynomial and SvF-method

The formulation seems to be more consistent with the physics of the phenomenon – regularization occurs at the level of functions that determine the solution, and not at the solution itself. See figure 3.

Errors (°C): σ^*_{cv}=2.22, rmsd*=1.82, Δ=0.83.

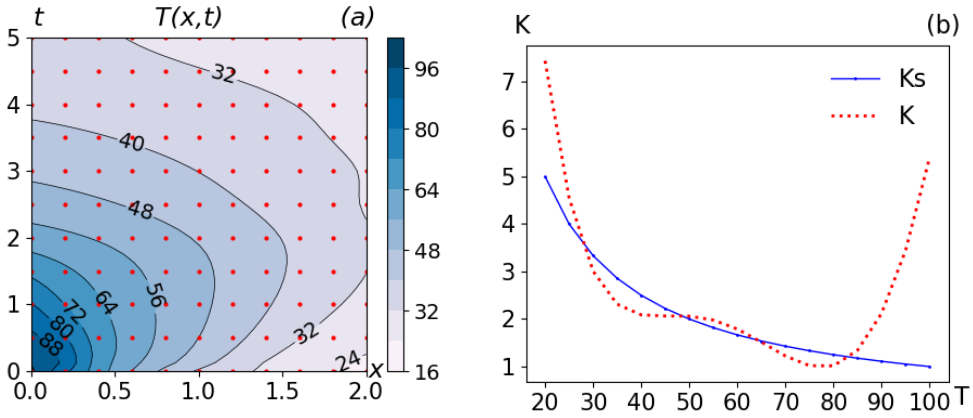


Figure 3. SvF solution of Problem 6.2. (a) contour lines of the solution *T(x,t)*; (b) thermal conductivity: *K* – solution, *Ks* – exact solution.

6.3. K(T) – Smooth-polyline and SvF-method

The setting differs from the previous one by using Smooth-polyline (the number of nodes is 16, ε=0.001) for approximation *K(T)*. See figure 4.

Errors (°C): σ^*_{cv}=2.20, rmsd*= 1.78, Δ=0.82.

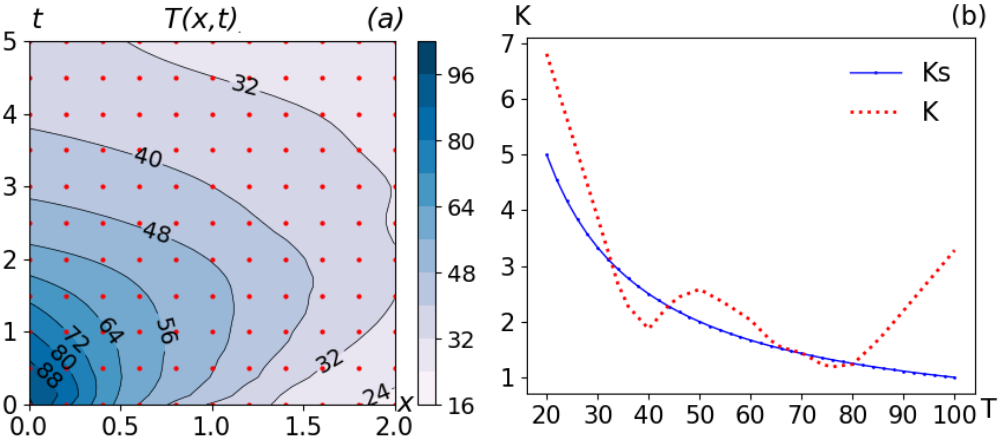


Figure 4. SvF solution of Problem 6.3. (a) contour lines of the solution *T(x,t)*; (b) thermal conductivity: *K* – solution, *Ks* – exact solution.

7. Discussion

Solution 5.1 describes the temperature dynamics relatively well (figure 2a). However, the lack of regularization causes the least squares method to try to repeat random errors in the solution, which leads to a "motley picture" and a far from accurate solution for the thermal conductivity (figure 2b).

The use of regularization enables to smooth the solution T(x,t) noticeably (see figure 2a and 3a) and increase the modeling accuracy. Pay attention to the incorrect behavior of the thermal conductivity near the right border of the graph in figures 2b and 3b. This is due to the small amount of experimental data in the region (lower left corner of the graph) where the temperature is close to 100 degrees.

Generally speaking, problems 5.2 and 5.3 lead to almost identical solutions. Thus, the use of an approximation method alternative to polynomials has been demonstrated. As expected, adding regularization terms leads to a more accurate (see Δ) modification of the model. In the technology used, this leads to a decrease in the σ^*_{cv} cross-validation error. It seems paradoxical that in this case the *rmsd** error grows: the more accurate the model, the greater its root mean square deviation from observations. However, it is easy to explain. First of all, rmsd* is within the error limits of the initial data σ_d. Second, the better the model, the closer it is to the exact solution, and for the exact solution $rmsd = \sigma_d$. Of course, if too large regularization penalties are chosen, the solution will be distorted so that *rmsd* will be greater than σ_d.

8. Conclusions

Regularization methods have a number of useful properties that make them efficient in solving inverse problems of experiments processing:

- the ability to process noisy data with significant errors and to assess how this affects the processing results,
- relative independence of numerical models grids from sets (grids) of measurements,
- the ability to use fewer measurements, as regularization methods have inherent properties of interpolation and extrapolation.

The problems (and their solution) considered in the article illustrate the effectiveness of the application of regularization methods and, in particular, the use of the balanced identification technology. As for computing resources, SvF technology is resource-intensive. This is justified as it is aimed at saving the researcher's time.

Acknowledgments

This research was carried out with the financial support of the Russian Scientific Foundation within the framework of the scientific project (grant) 22-11-00317. This work has been carried out using computing resources of the federal collective usage center Complex for Simulation and Data Processing for Megascience Facilities at NRC “Kurchatov Institute”.

References

[1] Tikhonov A N, Goncharsky A V, Stepanov V V and Yagola A G 1995 *Numerical methods for the solution of ill-posed problems* vol 328 (Springer Science & Business Media)

[2] Sokolov A V and Voloshinov V V 2020 Model selection by balanced identification: the interplay of optimization and distributed computing *Open Computer Science* **10** 283–95 https://doi.org/10.1515/comp-2020-0116

[3] Tikhonov A N 1980 On mathematical methods for automating the processing of observations. In: *Computational Mathematics Problems* (Moscow, Russia: Moscow State University Publishing House) pp 3-17

[4] Chen C and Mangasarian O L 1996 A class of smoothing functions for nonlinear and mixed complementarity problems *Comp. Opt. and App.* **5(2)** 97–138

[5] Zhou Z and Peng Y 2019 The locally Chen–Harker–Kanzow–Smale smoothing functions for mixed complementarity problems *J. of Global Opt* **74(1)** 169–93

[6] Samarskii A A 2001 *The theory of difference schemes* (New York: Marcel Dekker)