

# TAYLOR-MAP-BASED POLYNOMIAL APPROXIMATION OF NON-HOMOGENEOUS DYNAMICAL SYSTEMS

**Timofei Y. Gladyshev**

Saint Petersburg State University  
Saint Petersburg, Russian Federation  
st051995@student.spbu.ru

**Anna G. Golovkina**

Saint Petersburg State University  
Saint Petersburg, Russian Federation  
a.golovkina@spbu.ru

**Vladimir A. Kozynchenko**

Saint Petersburg State University  
Saint Petersburg, Russian Federation  
v.kozynchenko@spbu.ru

Article history:

Received 15.06.2026, Accepted 20.09.2026

## Abstract

The paper addresses the construction of a polynomial approximation of the flow of a dynamical system governed by autonomous ordinary differential equations with polynomial right-hand sides that contain non-homogeneous (constant) terms, in the case when the initial state is not close to the origin. The existing matrix Taylor map construction is restricted to homogeneous right-hand sides and to initial states in a neighborhood of zero.

The contribution of the paper is twofold. First, an algorithm for constructing the matrix Taylor map is developed that combines (i) expansion in deviations from the actual initial state,  $Y = X - X_0$ , and (ii) augmentation of the phase space by a constant auxiliary variable, which reduces the non-homogeneous problem to a homogeneous one; the Taylor map is then obtained from the matrizant of the associated linear system for phase moments. Second, the construction is validated numerically on a Van der Pol oscillator with constant bias: a fifth-order map reproduces the displaced limit cycle over  $t \in [0, 100]$  with an absolute error of  $10^{-4}$ – $10^{-3}$  after the transient, and expansion around  $X_0$  remains accurate in the regime where the expansion around zero degrades.

Unlike numerical integration, the construction yields an explicit parametric representation of the evolution operator whose coefficients are determined by the physical parameters of the system, and, unlike purely data-driven machine learning models, it exploits the governing equations. Since the coefficients enter linearly and have the same monomial structure as a dynamic polynomial regression model, the map can serve as a physically informed initial approximation for data-driven refinement;

this application is motivated by prior work and is not evaluated experimentally here.

## Key words

Taylor map, transfer map, polynomial regression, nonlinear dynamical systems, Van der Pol oscillator, Carleman linearization.

## 1 Introduction

The need to forecast the state of physical and technological processes over time arises in many fields, including vibration dynamics of industrial equipment, electrical oscillatory circuits, and charged particle beams. In many applications, it is sufficient to predict the process state at discrete, equally spaced time instants. This paper considers a dynamic regression model of the form

$$\mathcal{M} : X_{v,t_k} \rightarrow X_{v,t_{k+1}}, \quad (1)$$

where  $X_{v,t} = (x_{1,v,t}, \dots, x_{n,v,t})^\top$  is the state vector of the process at time  $t$ ,  $v$  is a parameter vector describing operating conditions, and  $t_0, \dots, t_N$  are equally spaced time moments, i.e.,  $t_{i+1} - t_i = t_{j+1} - t_j$  for all  $i, j = 0, \dots, N - 1$ .

The objective is to construct a one-step mapping that predicts the next state from the current one; in what follows, the current state  $X_{v,t_k}$  plays the role of the initial condition on each prediction step. In physical terms,  $\mathcal{M}$  is a discrete-time approximation of the flow (evolution operator) of the system over one time step.

The problem of constructing dynamic regression models of the form (1) is well studied. Among existing

approaches, one should note various machine learning models, including linear and polynomial regression, support vector machines, decision trees (e.g., CART), and neural network architectures such as feedforward networks (MLP, CNN) and recurrent networks (RNN, LSTM, GRU) [Jansson and Wahlberg, 1996; Herceg et al., 2019; Chen and Guestrin, 2016; Hornik et al., 1989; Rao et al., 1992; Abrevaya et al., 2018]; deep architectures have also been applied to the prediction of chaotic dynamics and extreme events [Gromov et al., 2024]. Ensemble methods, including random forests and gradient boosting, are also widely used [Chen and Guestrin, 2016]. These approaches require a sufficiently large training set, i.e., a sufficient number of state vectors  $X_{v,t_k}$  measured under different values of  $v$  and  $t_k$ . At the same time, apart from the measured trajectories, no additional information about the modeled process is required.

An alternative class of methods builds dynamic regression models on the basis of a differential description of the process, in particular, systems of ordinary differential equations [Kaheman et al., 2020; Brunton et al., 2016]. Such approaches typically involve two stages. The first stage addresses parameter identification, i.e., determining the coefficients of the differential system from available data. The second stage constructs the regression model by solving the identified system, usually by means of numerical integration.

The two-stage scheme is not the only option. Adaptive (on-line) identification methods estimate the parameters recursively along the measured trajectory, so that a separate pre-identification stage is not required. These include gradient and recursive least-squares estimators [Ljung, 1999; Ioannou and Sun, 1996], the speed-gradient approach [Fradkov, 2007], estimators based on dynamic regressor extension and mixing (DREM) [Aranovskiy et al., 2017], and filtering-based recursive identification schemes [Tsyganov and Tsyganova, 2025]; identification of dynamical models under time-varying disturbances and errors in variables is considered, e.g., in [Ivanov et al., 2024]. Adaptive estimators, however, act on the parameters of the differential model itself: they require a linearly parameterized model, typically a sufficiently exciting regressor (a condition relaxed, e.g., by DREM), and forecasting over a finite horizon still requires integration of the identified equations. The representation developed in the present paper is complementary rather than competing. Since the Taylor map is linear in its coefficients, it provides a linear regression form to which adaptive estimation laws can be applied directly, with the ODE-based map serving as the initial estimate. A quantitative comparison with adaptive estimators is beyond the scope of this paper and is planned for future work.

Solutions obtained by standard numerical solvers are not directly adaptable in parametric form: once computed for a fixed set of parameters, they cannot be efficiently refined using observed training vectors

$X_{v,t_k}$  without repeating the entire numerical integration. Therefore, to obtain a dynamic regression model that can be fine-tuned on data, one needs a representation of the solution that explicitly depends on adjustable parameters.

A natural choice is to represent the solution as a polynomial expansion in powers of the initial state vector  $X_0$ . This representation offers several important advantages:

- the expansion coefficients enter linearly, which allows efficient refinement by standard regression techniques;
- once constructed, the solution for a new initial condition can be evaluated directly, without repeated numerical integration;
- for a wide class of ordinary differential equations, the solution depends analytically on the initial vector, which guarantees the existence of such an expansion;
- the low-order terms describe linearized dynamics, whereas higher-order terms capture nonlinear interactions, yielding a structured representation of the system behavior.

Constructing the solution as a polynomial expansion in powers of the initial state vector thus provides a principled way to preserve the underlying dynamic structure while retaining tunable parameters. Such polynomial transfer maps originated in beam physics, where the motion of particles through accelerator elements is described by matrix maps acting on phase moments [Dragt, 2011; Andrianov, 2004]. The matrix formalism was introduced for autonomous homogeneous ODE systems [Andrianov, 2001; Andrianov, 1993] and later developed systematically for dynamical system identification [Golovkina et al., 2021a; Golovkina et al., 2022]. The corresponding procedure is based on the method of successive approximations and, importantly, requires the initial vector to lie in a neighborhood of the zero vector.

Taken together, the existing approaches leave an important gap. Machine learning models require large amounts of training data and do not explicitly exploit the differential structure of the process. Numerical ODE solvers, although physically grounded, do not provide a parametric representation that can be directly refined on observed trajectories. Finally, the polynomial expansion method based on successive approximations is restricted to homogeneous right-hand sides and to initial conditions near the origin.

The present study extends this polynomial expansion approach to autonomous systems with non-homogeneous right-hand sides and removes the requirement that the initial vector lie near zero. As a result, an approximate solution can be constructed for a substantially broader class of initial conditions and system structures, while retaining the dynamic interpretation of the underlying ODE system.

Since the approximate solution is obtained in explicit

polynomial form, its coefficient matrices have the same structure as the weight tensors of a dynamic polynomial regression model representing the mapping (1). In principle, they can therefore be used as a physically informed initial approximation before data-driven refinement. This is practically important because even polynomial models of modest degree and dimension involve a large number of parameters, and random initialization combined with limited training data often leads to poor convergence. [Golovkina et al., 2022] reports that identifying an approximate ODE-based solution from measured trajectories and using it to initialize polynomial regression improves optimization in that setting; [Ivanov et al., 2020] reports a similar benefit for polynomial neural networks. The present paper focuses on the analytical construction that makes such initialization possible for non-homogeneous systems and for initial states away from the origin; the numerical section validates the quality of the constructed Taylor map itself, not the subsequent training pipeline.

The remainder of the paper is organized as follows. Section 2 formulates the problem, describes the transition to deviation variables and the homogenization of a non-homogeneous polynomial ODE by phase-space augmentation, and discusses the physical meaning of these steps. Section 3 presents the algorithm for constructing the polynomial approximate solution (Taylor map). Section 4 reports numerical experiments on a Van der Pol benchmark with known coefficients. Section 5 concludes the paper.

## 2 Problem Statement

Consider a dynamic process modeled by a system of ordinary differential equations with a polynomial right-hand side

$$\frac{dX}{dt} = F(X) = \sum_{m=0}^N A^{1m}(t) X^{[m]}, \quad X(t_0) = X_0, \quad (2)$$

where  $X(t)$  is the phase state vector of dimension  $n$ ;  $X_0$  is the initial vector, which is not assumed to be close to zero;  $A^{1m}(t)$  are coefficient matrices of dimension  $n \times n^m$ ;  $A^{10}(t)$  is a vector of free terms (non-homogeneity), and  $X^{[0]} = 1$ . Here  $X^{[m]} = \underbrace{X * \dots * X}_{m \text{ times}}$

is the  $m$ -th Kronecker power of  $X$ , and  $*$  denotes the Kronecker product of matrices [Andrianov, 2001]. For generality, the coefficients are allowed to depend on time in Sections 2 and 3; in the autonomous case, which is the main subject of the paper and is considered in Section 4, they are constant.

**Full and reduced Kronecker powers.** The vector  $X^{[m]}$  of dimension  $n^m$  contains repeated monomials (e.g., for  $m = 3$  the monomial  $x^2y$  appears as  $xyx$ ,  $xyx$  and  $yxx$ ). Therefore, the coefficient matrices in (2) are not unique: the coefficient of a monomial can be distributed arbitrarily among its repeated positions without changing the vector field  $F(X)$ . In computations, we use the reduced Kronecker power, which contains each

distinct monomial of degree  $m$  exactly once, in lexicographic order, and has dimension  $\binom{n+m-1}{m}$ . The transition from the full to the reduced form consists in summing the columns of a coefficient matrix that correspond to identical monomials (and, in the equations for higher moments in Section 3, removing repeated rows). In the reduced form, the coefficient matrices are unique. The derivations below are written for full Kronecker powers, and the same symbols are used for reduced ones.

The standard Taylor map construction [Andrianov, 2001; Golovkina et al., 2022] applies to homogeneous systems ( $A^{10} = 0$ ) and expands the solution in powers of  $X_0$  about the origin, which requires  $\|X_0\|$  to be small. Both restrictions are removed below.

### 2.1 Transition to the Deviation Variable

Let us introduce the variable  $Y(t)$  describing the deviation of the trajectory from the initial point:

$$Y(t) = X(t) - X_0. \quad (3)$$

Since  $X_0$  is a constant,  $dY/dt = dX/dt$ . To rewrite the right-hand side of (2) in powers of  $Y$ , we use the Kronecker binomial expansion. By the mixed-product property of the Kronecker product,

$$(X_0 + Y)^{[m]} = \sum_{k=0}^m C^{mk}(X_0) Y^{[k]}, \quad (4)$$

$$C^{mk}(X_0) = \sum_{\sigma \in S_{m,k}} M_1^\sigma * \dots * M_m^\sigma,$$

where  $S_{m,k}$  is the set of binary sequences  $\sigma = (\sigma_1, \dots, \sigma_m)$  containing exactly  $k$  ones,  $M_j^\sigma = E$  if  $\sigma_j = 1$  and  $M_j^\sigma = X_0$  if  $\sigma_j = 0$ ,  $E$  is the identity matrix of order  $n$ , and  $Y^{[0]} = 1$ . For example,  $C^{21}(X_0) = X_0 * E + E * X_0$ ,  $C^{m0}(X_0) = X_0^{[m]}$ , and  $C^{mm} = E^{[m]}$ . Substituting (3) and (4) into (2), we obtain a system with respect to the new variable:

$$\frac{dY}{dt} = \sum_{k=1}^N P^{1k}(t) Y^{[k]} + B(t), \quad Y(t_0) = 0, \quad (5)$$

where

$$P^{1k}(t) = \sum_{m=k}^N A^{1m}(t) C^{mk}(X_0), \quad (6)$$

$$B(t) = \sum_{m=0}^N A^{1m}(t) X_0^{[m]} = F(X_0).$$

In particular,  $P^{11} = \partial F / \partial X|_{X=X_0}$  is the Jacobian matrix at the initial state. Note that the system (5) is non-homogeneous whenever  $X_0$  is not an equilibrium, i.e.,  $F(X_0) \neq 0$ , even if the original system is homogeneous ( $A^{10} = 0$ ). Thus, the transition to deviation variables and the treatment of free terms are two sides of the same problem, which is solved in the next subsection.

## 2.2 System Homogenization (Vector Augmentation)

To reduce the system (5) to a homogeneous form, we use the method of state space augmentation. Let us introduce an additional scalar component  $z$  satisfying  $dz/dt = 0$ , so that  $z(t) \equiv z(t_0)$ . We set

$$z(t) \equiv 1, \quad (7)$$

and represent the free term  $B(t)$  as  $B(t)z$ . With  $z \equiv 1$  this term coincides with  $B(t)$ , so no approximation is introduced; the symbol  $z$  is retained only to include the free term in the general polynomial structure and to track the order of the terms generated by it. Let us form an augmented vector  $\tilde{X}$  of dimension  $n + 1$ :

$$\tilde{X} = \begin{pmatrix} Y \\ z \end{pmatrix}. \quad (8)$$

Differentiating  $\tilde{X}$  by virtue of system (5) and the condition on  $z$ , we obtain:

$$\frac{d\tilde{X}}{dt} = \begin{pmatrix} \frac{dY}{dt} \\ \frac{dz}{dt} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^N P^{1k}(t)Y^{[k]} + B(t)z \\ 0 \end{pmatrix}, \quad (9)$$

where, by (7), the first  $n$  components coincide with the right-hand side of (5). Now the original problem can be rewritten as a homogeneous system of differential equations with respect to the vector  $\tilde{X}$ :

$$\frac{d\tilde{X}}{dt} = \sum_{k=1}^N \tilde{P}^{1k}(t)\tilde{X}^{[k]}, \quad (10)$$

where  $\tilde{P}^{1k}(t)$  are block coefficient matrices of the augmented system. The matrix  $\tilde{P}^{11}(t)$  is formed from the matrix  $P^{11}(t)$  and the vector  $B(t)$ :

$$\tilde{P}^{11}(t) = \begin{pmatrix} P^{11}(t) & B(t) \\ 0_{1 \times n} & 0 \end{pmatrix}.$$

Higher-order matrices  $\tilde{P}^{1k}(t)$  ( $k > 1$ ) are obtained from  $P^{1k}(t)$  by inserting zero columns at the positions of the monomials containing  $z$  and appending a zero row corresponding to the equation for  $z$ . The initial condition is

$$\tilde{X}(t_0) = \tilde{X}_0 = \begin{pmatrix} 0_{n \times 1} \\ 1 \end{pmatrix}. \quad (11)$$

Thus, the non-homogeneous Cauchy problem is reduced to a homogeneous system (10) in a higher-dimensional space.

## 2.3 Physical Interpretation

Each step of the construction has a transparent physical meaning. The vector  $B = F(X_0)$  in (6) is the phase velocity of the system at the initial state (for a mechanical system written in the first-order form, these are the velocities and the accelerations produced by the forces acting at  $X_0$ ), the matrix  $P^{11}$  describes the linearized

dynamics about the current state, and the matrices  $P^{1k}$ ,  $k \geq 2$ , describe local nonlinear couplings (anharmonic corrections).

The classical expansion around the origin is, in physical terms, a theory of small deviations from an equilibrium located at the origin, and its accuracy is governed by the amplitude of the motion. Expansion around the actual state  $X_0$  replaces this global smallness requirement by a local one: only the displacement  $Y = X(t) - X_0$  accumulated over the prediction step has to be small, and to leading order it equals  $F(X_0)(t - t_0)$ . Consequently, large-amplitude regimes, such as relaxation oscillations, limit cycles displaced from the origin by an external bias, and motions far from equilibrium, become accessible for a fixed expansion order, provided the time step is small compared with the characteristic time scale of the motion.

The auxiliary variable  $z$  has the meaning of a constant external action (a DC bias, a constant force or torque, a constant inflow) promoted to a dynamical variable. Since  $dz/dt = 0$ ,  $z$  is a first integral of the augmented system, the augmented phase space is foliated by the invariant hyperplanes  $z = \text{const}$ , and the physical system corresponds to the hyperplane  $z = 1$ . This is analogous to the extended phase space used in analytical mechanics to make non-autonomous systems autonomous. Since all components of  $\tilde{X}_0$  except  $z$  vanish, the resulting Taylor map is a power series in  $z$  (see (26) below), i.e., a perturbation expansion in the amplitude of the effective forcing  $Bz$ , evaluated at  $z = 1$ ; this forcing is precisely the quantity that drives the system away from its initial state.

Finally, the construction yields a polynomial approximation of the flow of the physical system over the time step. As noted in the Introduction, such transfer maps are a standard tool of beam physics [Dragt, 2011; Andrianov, 2004], and the embedding of a nonlinear system into a linear system for phase moments used in Section 3 is the Carleman linearization [Carleman, 1932]. Every coefficient of the map is determined by the coefficients of the governing equations (e.g., stiffness, damping, bias) through the explicit formulas of Section 3, which makes the model interpretable and allows the physical parameters to be traced through subsequent data-driven refinement.

## 3 Construction of the Polynomial Approximate Solution

We now construct an approximate solution of the homogeneous augmented system (10) obtained in Section 2. In applications where the coefficient matrices  $\tilde{P}^{1k}(t)$  are not known a priori, they can be estimated from measured trajectories using the reconstruction procedure of [Golovkina et al., 2022] or the adaptive methods mentioned in Section 1; the numerical experiments in Section 4 use analytically specified matrices instead. Throughout this section,  $\tilde{X}$  and  $\tilde{P}^{1k}$  refer to the augmented system (10); the matrices  $A^{1m}$  of the original system (2) and  $P^{1k}$  of the deviation system (5) enter only through the construction of  $\tilde{P}^{1k}$  in Section 2.2.

### 3.1 Increasing the Dimensionality of the System of Differential Equations

To construct an approximate solution of the system (10), let us increase its dimensionality. To do this, we introduce the vector of phase moments up to order  $K$ :  $\tilde{X}^K = [\tilde{X}^{[1]}, \tilde{X}^{[2]}, \dots, \tilde{X}^{[K]}]^T$  and calculate the derivatives for all vectors  $\tilde{X}^{[i]}, i = \overline{1, K}$ . The derivative of the first-order vector is represented by the system (10). Let us calculate the derivative of the vector of second-order phase moments:

$$\begin{aligned} \frac{d\tilde{X}^{[2]}}{dt} &= \frac{d(\tilde{X} * \tilde{X})}{dt} = \frac{d\tilde{X}}{dt} * \tilde{X} + \tilde{X} * \frac{d\tilde{X}}{dt} \\ &= \sum_{k=1}^N \tilde{P}^{1k}(t) \tilde{X}^{[k]} * \tilde{X} + \tilde{X} * \sum_{k=1}^N \tilde{P}^{1k}(t) \tilde{X}^{[k]} \\ &= \sum_{k=1}^N \tilde{P}^{1k}(t) \tilde{X}^{[k]} * (E\tilde{X}) + \sum_{k=1}^N (E\tilde{X}) * \tilde{P}^{1k}(t) \tilde{X}^{[k]} \\ &= \sum_{k=1}^N \left( \tilde{P}^{1k}(t) * E \right) \tilde{X}^{[k+1]} \\ &\quad + \sum_{k=1}^N \left( E * \tilde{P}^{1k}(t) \right) \tilde{X}^{[k+1]} \\ &= \sum_{k=2}^{N+1} \tilde{P}^{2k}(t) \tilde{X}^{[k]}, \end{aligned} \quad (12)$$

where

$$\tilde{P}^{2k}(t) = \tilde{P}^{1, k-1}(t) * E + E * \tilde{P}^{1, k-1}(t).$$

Here  $E$  is the identity matrix of the dimension of the vector  $\tilde{X}$ . When reduced Kronecker powers are used (Section 2), the columns of  $\tilde{P}^{2k}$  corresponding to identical monomials of  $\tilde{X}^{[k]}$  are summed and the repeated rows are removed. By induction, one can show the general formula for the derivative of a vector of arbitrary order  $i$ :

$$\frac{d\tilde{X}^{[i]}}{dt} = \sum_{k=i}^{N+i-1} \tilde{P}^{ik}(t) \tilde{X}^{[k]}, \quad (13)$$

where

$$\tilde{P}^{ik}(t) = \tilde{P}^{1, k-i+1}(t) * E^{[i-1]} + E * \tilde{P}^{i-1, k-1}(t). \quad (14)$$

From the representation (13), it follows that the equation (10) can be associated with an equation in the space of phase moments up to order  $K$ :

$$\frac{d\tilde{X}^K}{dt} = \tilde{H}^K \tilde{X}^{K+N-1}, \quad (15)$$

where

$$\begin{aligned} \tilde{P}^K &= \begin{pmatrix} \tilde{P}^{11} & \tilde{P}^{12} & \tilde{P}^{13} & \dots & \tilde{P}^{1K} \\ \mathbb{O} & \tilde{P}^{22} & \tilde{P}^{23} & \dots & \tilde{P}^{2K} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & \mathbb{O} & \tilde{P}^{KK} \end{pmatrix}, \\ \tilde{G}^K &= \begin{pmatrix} \tilde{P}^{1, K+1} & \dots & \tilde{P}^{1, K+N-1} \\ \vdots & \ddots & \vdots \\ \tilde{P}^{K, K+1} & \dots & \tilde{P}^{K, K+N-1} \end{pmatrix}, \end{aligned} \quad (16)$$

Here, the matrices  $\tilde{P}^{ik}(t)$  are calculated using recurrent formulas (14), and  $\tilde{P}^{1k}(t) = \mathbb{O}$  for  $k > N$ . In particular, the first block row of  $\tilde{P}^K$  reproduces the right-hand side of (10).

### 3.2 Linearized System of Increased Dimensionality and Its Solution

Consider the linear part of the system (15), discarding the terms containing moments of orders higher than  $K$ :

$$\frac{d\tilde{X}^K}{dt} = \tilde{P}^K \tilde{X}^K. \quad (17)$$

The system (17) is linear in form, so the solution of the Cauchy problem  $\tilde{X}^K(t_0) = \tilde{X}_0^K$  can be written using a nonlinear matrizant:

$$\tilde{X}^K(t) = \tilde{R}^K(t, t_0) \tilde{X}_0^K, \quad (18)$$

where  $\tilde{R}^K(t_0, t_0) = E^K$ . The matrizant  $\tilde{R}^K$  has an upper-triangular block structure similar to  $\tilde{P}^K$ :

$$\tilde{R}^K = \begin{pmatrix} \tilde{R}^{11} & \tilde{R}^{12} & \tilde{R}^{13} & \dots & \tilde{R}^{1K} \\ \mathbb{O} & \tilde{R}^{22} & \tilde{R}^{23} & \dots & \tilde{R}^{2K} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & \mathbb{O} & \tilde{R}^{KK} \end{pmatrix}.$$

To find the elements of  $\tilde{R}^K$ , we represent the system (17) as a cascade of  $K$  subsystems:

$$\begin{cases} \frac{d\tilde{X}}{dt} = \tilde{P}^{11}(t) \tilde{X} + \tilde{P}^{12}(t) \tilde{X}^{[2]} \\ \quad + \dots + \tilde{P}^{1K}(t) \tilde{X}^{[K]} \\ \frac{d\tilde{X}^{[2]}}{dt} = \tilde{P}^{22}(t) \tilde{X}^{[2]} + \dots + \tilde{P}^{2K}(t) \tilde{X}^{[K]} \\ \dots \\ \frac{d\tilde{X}^{[K-1]}}{dt} = \tilde{P}^{K-1, K-1}(t) \tilde{X}^{[K-1]} + \tilde{P}^{K-1, K}(t) \tilde{X}^{[K]} \\ \frac{d\tilde{X}^{[K]}}{dt} = \tilde{P}^{KK}(t) \tilde{X}^{[K]} \end{cases} \quad (19)$$

For the diagonal elements, consider the linearized first equation  $\frac{d\tilde{X}}{dt} = \tilde{P}^{11}(t) \tilde{X}$ . Its solution has the form  $\tilde{X}(t, t_0) = \tilde{R}^{11}(t, t_0) \tilde{X}_0$ , where in the autonomous case (constant matrix  $\tilde{P}^{11}$ ) the matrizant is calculated as:

$$\tilde{R}^{11}(t, t_0) = e^{\tilde{P}^{11}(t-t_0)}. \quad (20)$$

Raising this solution to Kronecker powers, we obtain the fundamental property of diagonal blocks:

$$\tilde{R}^{ii}(t, t_0) = \left( \tilde{R}^{11}(t, t_0) \right)^{[i]}. \quad (21)$$

Non-diagonal elements are found using the backward substitution method (“bottom-up”). Solving the penultimate subsystem as a linear non-homogeneous one using Cauchy’s formula, we obtain the coupling matrizant:

$$\tilde{R}^{K-1, K}(t, t_0) = \int_{t_0}^t \tilde{R}^{K-1, K-1}(t, \tau) \tilde{P}^{K-1, K}(\tau) \tilde{R}^{KK}(\tau, t_0) d\tau. \quad (22)$$

Generalizing by induction, we obtain the final recurrence relations for the matrices  $\tilde{R}^{ij}(t, t_0)$  necessary to reconstruct the solution:

$$\tilde{R}^{1j}(t, t_0) = \sum_{l=2}^j \int_{t_0}^t \tilde{R}^{11}(t, \tau) \tilde{P}^{1l}(\tau) \tilde{R}^{lj}(\tau, t_0) d\tau, \quad (23)$$

$$\begin{aligned} \tilde{R}^{ij}(t, t_0) &= \sum_{l=i+1}^j \int_{t_0}^t \tilde{R}^{ii}(t, \tau) \tilde{P}^{il}(\tau) \tilde{R}^{lj}(\tau, t_0) d\tau, \\ 2 \leq i < j \leq K. \end{aligned} \quad (24)$$

**Remark on truncation and uniqueness.** Owing to the upper block-triangular structure of  $\tilde{P}^K$ , the block  $\tilde{R}^{1k}$  depends only on the blocks  $\tilde{P}^{ij}$  with  $i \leq j \leq k$ . Hence, discarding the moments of orders higher than  $K$  in (17) does not affect  $\tilde{R}^{1k}$  for  $k \leq K$ , and  $\tilde{R}^{1k}(t, t_0) \tilde{X}_0^{[k]}$  coincides with the  $k$ -th order term of the Taylor expansion of the exact flow of (10) in  $\tilde{X}_0$ . This term is unique as a polynomial function. When full Kronecker powers are used, the matrices  $\tilde{R}^{1k}$  themselves depend on how the coefficients are distributed among identical monomials, but their action on Kronecker powers does not; in the reduced form, the matrices are unique.

### 3.3 Back Substitution and Solution Recovery

After calculating the blocks of the matrizant  $\tilde{R}^{1k}(t, t_0)$  (where  $k = \overline{1, K}$ ) using the recurrence relations (23), (24), it is necessary to perform an inverse transition from the variables of the augmented phase space to the original variables  $X(t)$ .

#### Calculation of the deviation vector in the augmented space

From the system of equations (19), we are interested only in the first equation describing the dynamics of the vector  $\tilde{X}(t)$  (the first moment), since it is the one containing information about the physical variables:

$$\tilde{X}(t) \approx \sum_{k=1}^K \tilde{R}^{1k}(t, t_0) \tilde{X}_0^{[k]}. \quad (25)$$

Here, the summation takes into account the contributions of the terms of all orders up to  $K$  to the first-moment solution.

#### Structure of the initial vector

Recall the structure of the augmented initial vector  $\tilde{X}_0$  defined in (11). Since  $Y(t_0) = 0$ , the vector has the form:

$$\tilde{X}_0 = \begin{pmatrix} 0_{n \times 1} \\ z \end{pmatrix},$$

where  $0_{n \times 1}$  is a zero vector of dimension  $n$  and  $z = 1$  by (7); the symbol  $z$  is kept in the formulas below to show the order of each term. Let us consider the structure of the Kronecker power  $\tilde{X}_0^{[k]}$ . This is a vector of dimension  $(n+1)^k$  (or  $\binom{n+k}{k}$  in the reduced form). Since in the lexicographic order the last element of the resulting vector is formed as the product of the last elements of the original vectors, and all other combinations in  $\tilde{X}_0$  contain at least one zero, the vector  $\tilde{X}_0^{[k]}$  turns out to be sparse (in both the full and the reduced form, because  $z$  is the last variable):

$$\tilde{X}_0^{[k]} = \underbrace{\tilde{X}_0 * \dots * \tilde{X}_0}_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ z^k \end{pmatrix}. \quad (26)$$

It contains zeros in all positions except the very last one. Consequently, in the expression  $\tilde{R}^{1k}(t, t_0) \tilde{X}_0^{[k]}$ , we are interested only in the last column of the block  $\tilde{R}^{1k}$ , which significantly reduces the amount of computations.

#### Projection onto the original space

By construction, the vector  $\tilde{X}(t)$  consists of two parts: the desired deviation  $Y(t)$  and the auxiliary constant  $z$ :

$$\tilde{X}(t) = \begin{pmatrix} Y(t) \\ z \end{pmatrix}.$$

To extract the upper part of the vector (of dimension  $n$ ) and discard the lower scalar part, a projection matrix (selector)  $S$  of dimension  $n \times (n+1)$  is used:

$$S = (E \ 0_{n \times 1}),$$

where  $E$  is the identity matrix of order  $n$ . Indeed,

$$S \tilde{X}(t) = (E \ 0_{n \times 1}) \begin{pmatrix} Y(t) \\ z \end{pmatrix} = E \cdot Y(t) + 0_{n \times 1} \cdot z = Y(t).$$

#### Final solution formula

Substituting (25) into the expression for  $X(t) = X_0 + Y(t)$  and applying the selector  $S$ , we obtain:

$$X(t) \approx X_0 + S \sum_{k=1}^K \tilde{R}^{1k}(t, t_0) \tilde{X}_0^{[k]}. \quad (27)$$

Taking into account property (26), the formula can be written in a computationally efficient form. Let  $\{\tilde{R}^{1k}(t, t_0)\}_{last}$  denote the last column of the matrix  $\tilde{R}^{1k}$ . Then the final solution is:

$$X(t) \approx X_0 + \sum_{k=1}^K \left( \{\tilde{R}^{1k}(t, t_0)\}_{last} \cdot z^k \right)_{1..n}, \quad (28)$$

where the index  $(\dots)_{1..n}$  denotes taking the first  $n$  elements (upper part) of the obtained column vector, which is equivalent to the action of the selector  $S$ .

#### 4 Numerical Results

This section validates the proposed construction on a benchmark with known ODE coefficients. The experiments assess (i) multi-step forecasting accuracy for a non-homogeneous system and (ii) the effect of centering the expansion at  $X_0$  rather than at the origin. Coefficient identification from data and regression training are left for future work.

As a benchmark, we consider the Van der Pol oscillator, a classical model of self-sustained oscillations originally introduced for triode circuits [Cartwright, 1952], which remains a standard model of self-excited electrical and mechanical systems (see, e.g., [Shriethar, 2025] for a recent analysis of a mechanical system with Van der Pol-type dynamics). We consider the oscillator with a constant external action (bias):

$$\begin{cases} \dot{x} = y + c_1, \\ \dot{y} = -x + y(1 - x^2) + c_2, \end{cases} \quad (29)$$

where  $c_1$  and  $c_2$  are constants determining the magnitude of the non-homogeneity. Physically, they model a constant external action, such as a DC bias in the circuit or a constant external force, which shifts the equilibrium and the limit cycle away from the origin. This makes the system a natural test case for both components of the construction: the constant terms require homogenization, and the displaced limit cycle provides states far from the origin.

To apply the Taylor map method, the non-homogeneous system (29) is brought to a homogeneous form by augmenting the phase space with the auxiliary variable  $z$ ,  $\dot{z} = 0$ ,  $z(t_0) = 1$  (Section 2.2):

$$\begin{cases} \dot{x} = y + c_1 z, \\ \dot{y} = -x + y - x^2 y + c_2 z, \\ \dot{z} = 0. \end{cases} \quad (30)$$

Let us denote the augmented state vector as  $X_{aug} = (x, y, z)^T$ . Then the right-hand side of the system (30) can be represented in matrix form up to the third order of nonlinearity:

$$\frac{dX_{aug}}{dt} = P^{11} X_{aug} + P^{12} X_{aug}^{[2]} + P^{13} X_{aug}^{[3]}, \quad (31)$$

where the coefficient matrices  $P^{11}$ ,  $P^{12}$ ,  $P^{13}$  are fully determined by the structure of the equations. In our computational experiment, the bias parameters were set to  $c_1 = 0.1$  and  $c_2 = 1.0$ . The matrices are written in the reduced form (Section 2): the columns of  $P^{12}$  correspond to the monomials  $x^2, xy, xz, y^2, yz, z^2$ , and the columns of  $P^{13}$  correspond to  $x^3, x^2 y, x^2 z, xy^2, xyz, xz^2, y^3, y^2 z, yz^2, z^3$ . The matrix of the linear part  $P^{11}$  has dimensions  $3 \times 3$ , the matrix of quadratic terms  $P^{12}$  of dimensions  $3 \times 6$  is zero in this case, and the matrix of cubic terms  $P^{13}$  of dimensions  $3 \times 10$  contains a single non-zero element  $P_{2,2}^{13} = -1$  corresponding to the monomial  $x^2 y$ :

$$\begin{aligned} P^{11} &= \begin{pmatrix} 0 & 1 & 0.1 \\ -1 & 1 & 1.0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P^{12} = \mathbf{0}_{3 \times 6}, \\ P^{13} &= \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ 0 & 0 & \dots & 0 \end{pmatrix}. \end{aligned} \quad (32)$$

In the reduced form, this representation is unique. If full Kronecker powers were used ( $P^{13}$  of dimensions  $3 \times 27$ ), the coefficient  $-1$  could be distributed arbitrarily among the positions of  $xxxy$ ,  $xyxy$  and  $yxyx$  with the total sum equal to  $-1$ . Such a choice changes the entries of the matrices but neither the vector field nor, by the remark in Section 3.2, the Taylor map as a function of the initial state.

Using the matrices  $P^{1k}$ , a fifth-order matrix Taylor map (a discrete dynamic polynomial model with the same monomial structure as the regression model (1)) was constructed. In this experiment, the expansion is performed about the origin of the augmented space, i.e., only the augmentation step of Section 2.2 is used, and the map is a fixed polynomial of  $(x, y, z)$ . The map was applied recursively with time step  $\Delta t = 0.01$  to emulate multi-step forecasting of model (1); no parameter fitting on trajectories was performed. The initial conditions for the simulation were set as  $x_0 = -2.0, y_0 = 4.0$  (respectively,  $X_{aug}(0) = (-2.0, 4.0, 1.0)^T$ ). The forecasting was carried out over the time interval  $t \in [0, 100]$ . To verify the results, numerical integration of the system (29) was performed in parallel using the standard solver LSODA [LSODA, 2005].

The simulation results are presented in Fig. 1. As can be seen from the phase portrait, the constant external action shifts the limit cycle relative to the origin. The trajectory generated by the Taylor map visually coincides with the reference solution over the entire prediction interval.

A quantitative assessment of the accuracy is presented in Fig. 2. The absolute error plots demonstrate that the maximum deviation of the predicted trajectory from the reference one during the transient does not exceed 0.035, and after reaching the limit cycle, the error stabilizes in

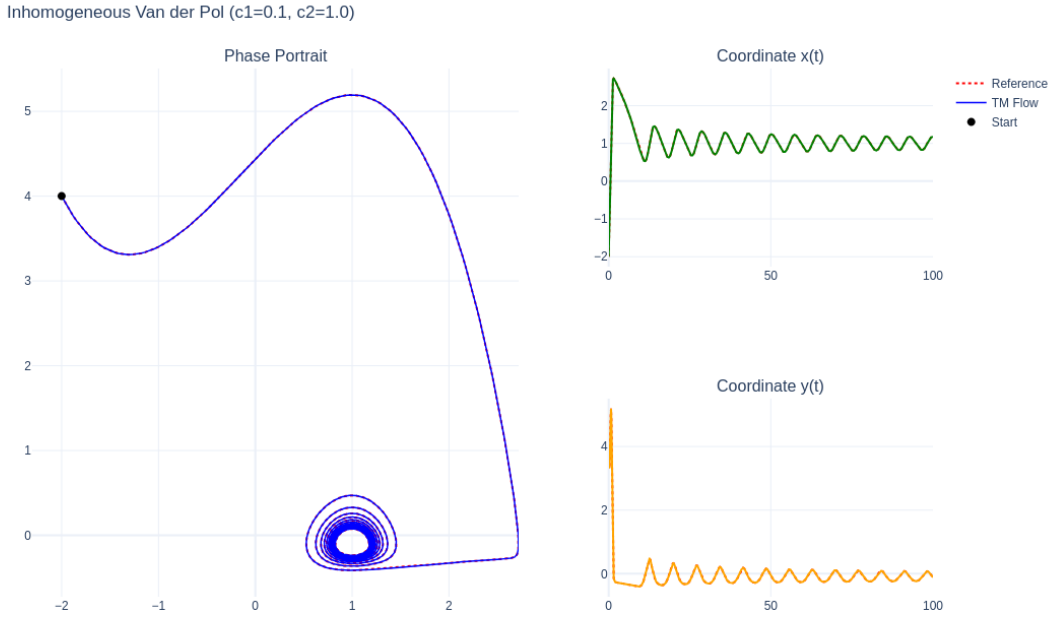


Figure 1. Comparison of the dynamics of a non-homogeneous Van der Pol oscillator ( $c_1 = 0.1, c_2 = 1.0$ ) obtained using the constructed Taylor map (TM Flow, blue line) and the reference numerical solution (Reference, dashed line). Left: phase portrait. Right: time sweep of the coordinates  $x(t)$  and  $y(t)$ .

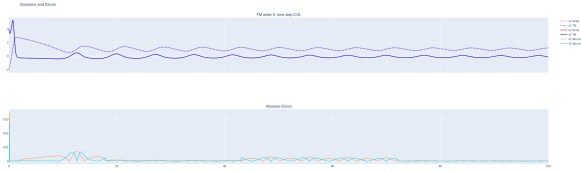


Figure 2. Dynamics of the absolute error  $|x_{TM} - x_{lsoda}|$  and  $|y_{TM} - y_{lsoda}|$  during multi-step forecasting with a step  $\Delta t = 0.01$  and model order equal to 5.

the range from  $10^{-4}$  to  $10^{-3}$ . This result confirms that the phase space augmentation technique reproduces non-homogeneous dynamics with high accuracy over long time intervals.

To further investigate the capabilities of the proposed approach, a second computational experiment was conducted to evaluate the effect of the expansion point on the accuracy of the Taylor map. This is particularly relevant when the initial state is located far from the origin.

In the standard method of successive approximations proposed in [Golovkina et al., 2022], the series expansion is performed around the zero point. By introducing the deviation variable  $Y = X - X_0$  (Section 2.1), the expansion can instead be centered at the initial point  $X_0$ . As shown by (6), this shift introduces the constant term  $B = F(X_0)$ , transforming the system into a non-homogeneous one, which is resolved by the same phase space augmentation technique ( $z = 1$ ) as in the first ex-

periment. Figure 3 illustrates the phase trajectories and coordinate dynamics for the reference numerical solution alongside two Taylor maps: the standard one expanded at zero [Golovkina et al., 2022] and the proposed one expanded at the initial point  $X_0$ .

A quantitative comparison of the absolute errors for both approaches is shown in Fig. 4. When the trajectory starts far from the origin, the expansion around zero [Golovkina et al., 2022] does not maintain accuracy: the two maps begin to diverge noticeably at  $t \approx 7$ . By  $t = 15$ , the absolute error of the map expanded around the initial point remains of the order of  $10^{-3}$ , whereas the error of the expansion around zero reaches values between 1 and 2.

This demonstrates that shifting the expansion point to  $X_0$  significantly improves the accuracy and stability of the constructed Taylor map when the initial state is far from the coordinate origin. In physical terms (Section 2.3), the expansion around zero treats the motion as a deviation from the origin, and its truncation error grows with the amplitude of the motion, whereas the expansion around the current state only requires the displacement over one step to be small.

## 5 Conclusion

In this paper, we extended the Taylor-map construction for dynamic polynomial models to autonomous systems with non-homogeneous polynomial right-hand sides and to initial conditions away from the origin. The extension

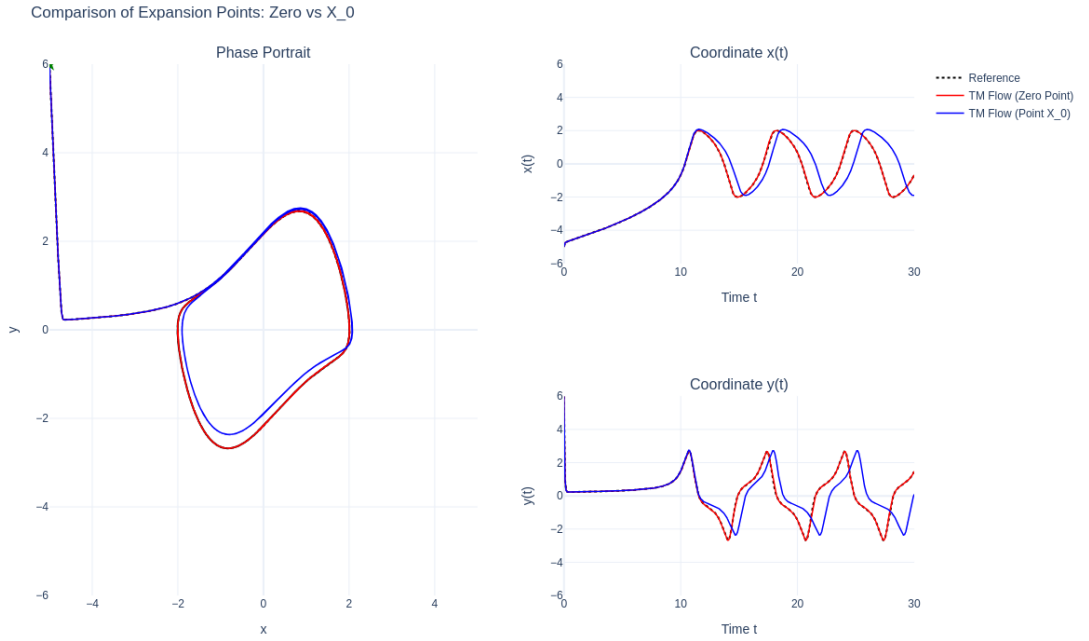


Figure 3. Comparison of the expansion points for the Van der Pol oscillator: the reference numerical solution (Reference, black dashed line), the Taylor map expanded around zero [Golovkina et al., 2022] (TM Flow (Zero Point), red line), and the proposed Taylor map expanded around the initial point (TM Flow (Point  $X_0$ ), blue line). Left: phase portrait. Right: time sweep of the coordinates  $x(t)$  and  $y(t)$ .

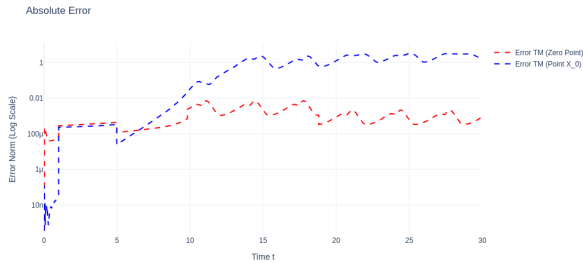


Figure 4. Dynamics of the absolute error of the Taylor map expanded around zero [Golovkina et al., 2022] and of the proposed Taylor map expanded around the initial point  $X_0$  (logarithmic scale).

uses expansion in deviations  $Y = X - X_0$  together with phase-space augmentation by an auxiliary constant variable, which homogenizes the system while keeping the expansion centered at the actual initial state.

From the physical point of view, the constructed map is a polynomial approximation of the flow of the system over the time step, of the same kind as the transfer maps used in beam physics. The constant term of the deviation system is the phase velocity at the initial state, its linear part is the Jacobian at this state, and the auxiliary variable represents a constant external action as a first integral of the augmented system. Expansion around the current state replaces the requirement of small amplitude of motion by the requirement of small displacement over one step, which makes large-amplitude and biased

regimes accessible for a fixed expansion order.

The algorithm produces an explicit polynomial approximate solution whose coefficient matrices have the same monomial structure as a dynamic polynomial regression model (1). This structural correspondence is the basis for using the construction as a physically informed initial approximation in data-driven refinement, including adaptive (on-line) estimation; prior studies suggest benefits of such initialization for polynomial regression and polynomial neural networks, but the present work does not experimentally test training or identification from noisy data.

Numerical validation was carried out on a non-homogeneous Van der Pol oscillator with known coefficients. In the first experiment, a fifth-order Taylor map reproduced the displaced limit cycle over  $t \in [0, 100]$  with  $\Delta t = 0.01$ ; after the initial transient, the absolute forecasting error stabilized between  $10^{-4}$  and  $10^{-3}$ . In the second experiment, expansion around  $X_0$  remained accurate, while the standard expansion around zero diverged when the trajectory started far from the origin. These results confirm the accuracy of the extended construction for non-homogeneous dynamics and for non-zero initial states.

The main limitations are that only one benchmark system was considered, the ODE coefficients were specified analytically rather than identified from measurements, and the regression initialization workflow was not compared with random initialization, purely data-driven models, or adaptive estimators. Future work will address

identification from irregular time series, fine-tuning of the sparse coefficient matrices on noisy experimental data, comparison with adaptive identification methods, and application to industrial dynamic processes.

## 6 Acknowledgements

The research was carried out at Saint Petersburg University, project no. 121061000159-6.

## References

- Abrevaya, G., Rish, I., Aravkin, A. Y., Cecchi, G., Kozloski, J., Polosecki, P., Zheng, P., Dawson, S. P., Rhee, J., and Cox, D. (2018). Learning nonlinear brain dynamics: van der Pol Meets LSTM. *bioRxiv* 330548.
- Andrianov, S. (1993). A matrix representation of the Lie transformation. In *Proceedings of the Abstracts of the International Congress on Computer Systems and Applied Mathematics*, St. Petersburg, **14**, pp. 19–23.
- Andrianov, S. (2001). Symbolic computation of approximate symmetries for ordinary differential equations. *Mathematics and Computers in Simulation*, **57** (3–5), pp. 147–154.
- Andrianov, S. N. (2004). *Dynamical modeling of control systems for particle beams*. Publishing house of Saint Petersburg State University, St. Petersburg.
- Aranovskiy, S., Bobtsov, A., Ortega, R., and Pyrkin, A. (2017). Performance enhancement of parameter estimators via dynamic regressor extension and mixing. *IEEE Transactions on Automatic Control*, **62** (7), pp. 3546–3550.
- Blondel, M., Ishihata, M., Fujino, A., and Ueda, N. (2016). Polynomial networks and factorization machines: New insights and efficient training algorithms. In *Proceedings of ICML 2016*, New York City, pp. 850–858.
- Blondel, M., Niculae, V., Otsuka, T., and Ueda, N. (2017). Multi-output polynomial networks and factorization machines. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*, Long Beach, pp. 3351–3361.
- Brunton, S. L., Joshua, L. P., and Kutz, N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, **113** (15), pp. 3932–3937.
- Carleman, T. (1932). Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires. *Acta Mathematica*, **59**, pp. 63–87.
- Cartwright, M. L. (1952). Van der Pol's equation for relaxation oscillations. In *Contributions to the Theory of Nonlinear Oscillations II*, Princeton University Press, Princeton, pp. 3–18.
- Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16)*, San Francisco, pp. 785–794.
- Dette, H. (1995). Optimal designs for identifying the degree of a polynomial regression. *Annals of Statistics*, **23** (4), pp. 1248–1266.
- Dragt, A. (2011). *Lie methods for nonlinear dynamics with applications to accelerator physics*. inspirehep.net/record/955313/files/TOC28Nov2011.pdf.
- Fradkov, A. L. (2007). *Cybernetical Physics: From Control of Chaos to Quantum Control*. Springer, Berlin, Heidelberg.
- Golovkina, A., Kozynchenko, V., and Kulabukhova, N. (2021). Reconstruction and identification of dynamical systems based on taylor maps. In *Computational Science and Its Applications — 21st International Conference, Proceedings, Pt VIII*, Springer Nature, Switzerland, pp. 360–369.
- Golovkina, A. G., Kozynchenko, V. A., and Kulabukhova, N. V. (2021). Reconstruction of ordinary differential equations from irregularly distributed time-series data. In *Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2021)*, Dubna, Russia, **3041**, pp. 342–347.
- Golovkina, A. G., and Ganaeva, D. D. (2022). Method of nonlinear dynamical systems reconstruction based on time series data. *Control processes and Stability*, **9** (1), pp. 197–201.
- Golovkina, A. G., Kozynchenko, V. A., and Klimenko, I. S. (2022). The method of successive approximations for constructing a model of dynamic polynomial regression. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, **18** (4), pp. 00–00.
- Gromov, N., Smirnov, L., and Levanova, T. (2024). Prediction of extreme events and chaotic dynamics using WaveNet. *Cybernetics and Physics*, **13** (1), pp. 20–31. doi:10.35470/2226-4116-2024-13-1-20-31.
- Herceg, S., Ujevic, Z., and Bolf, A. N. (2019). Development of soft sensors for isomerization process based on support vector machine regression and dynamic polynomial models. *Chemical Engineering Research and Design*, **149**, pp. 95–103.
- Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, pp. 359–366.
- Ioannou, P. A., and Sun, J. (1996). *Robust Adaptive Control*. Prentice Hall, Upper Saddle River, NJ.
- Ivanov, A., Golovkina, A., and Iben, U. (2020). Polynomial neural networks and taylor maps for dynamical systems simulation and learning. In *24th European Conference on Artificial Intelligence (PAIS 2020 Proceedings)*, IOS Press, pp. 1230–1237.
- Ivanov, D., Zhdanov, A., and Sandler, I. (2024). Identification of dynamical systems with time-varying disturbances and fractional errors-in-variables. Application example to a drone model. *Cybernetics and Physics*,

- 13**(1), pp. 49–56. doi:10.35470/2226-4116-2024-13-1-49-56.
- Jansson, M., and Wahlberg, B. (1996). A linear regression approach to state-space subspace system identification. *Signal Processing*, **2**, pp. 103–129.
- Kaheman, K., Kutz, J. N., and Brunton, S. L. (2020). SINDy-PI: a robust algorithm for parallel implicit sparse identification of nonlinear dynamics. *Proceedings of the Royal Society A*, **476**(2242), Art. N. 20200279.
- Kim, B., Ko, C.-Y., and Wong, N. (2018). Tensor Network Subspace Identification of Polynomial State Space Models. *Automatica*, **95**, pp. 187–196.
- Klimenko, I. S. (2022). Implementation of matrix map method for solving a system of ordinary equations. *Control processes and Stability*, **9**(1), pp. 53–57.
- Ljung, L. (1999). *System Identification: Theory for the User*, 2nd ed. Prentice Hall, Upper Saddle River, NJ.
- LSODA (2005). Ordinary differential equation solver for stiff or non-stiff system. <http://www.nea.fr/abs/html/uscd1227.html>.
- Rao, S., Sethuraman, S., and Ramamurthi, V. (1992). A recurrent neural network for nonlinear time series prediction: a comparative study. In *IEEE 1992 Workshop on Neural Networks for Signal Processing (NNSP'92)*, Helsingoer, pp. 531–539.
- Shriethar, N. (2025). Limit cycles and period-doubling bifurcations in a van der Pol-type triple pendulum with spring coupling. *Cybernetics and Physics*, **14**(2), pp. 174–190. doi:10.35470/2226-4116-2025-14-2-174-190.
- Tsyganov, A., and Tsyganova, J. (2025). Parameter identification of discrete-time linear stochastic systems based on decentralized square-root information filtering. *Cybernetics and Physics*, **14**(2), pp. 191–199. doi:10.35470/2226-4116-2025-14-2-191-199.