Preprints of the 2012 20th Mediterranean Conference on Control & Automation (MED), Barcelona, Spain, July 3-6, 2012

Initial Conditions Optimization of Nonlinear Dynamic Systems with Applications to Output Identification and Control

Josip Kasac, Vladimir Milic, Branko Novakovic, Dubravko Majetic and Danko Brezak

Abstract— **The paper presents a gradient-based algorithm for initial conditions optimization of nonlinear multivariable systems with boundary and state vectors constraints. The algorithm has a backward-in-time recurrent structure similar to the backpropagation-through-time (BPTT) algorithm, which is mostly used as a learning algorithm for dynamic neural networks. It is shown that dynamic parameter optimization problem can be formulated as the initial conditions optimization problem. Further, it is shown that output parameter identification and output controller design problems can be formulated as dynamic parameter optimization problem. The effectiveness of the proposed algorithm is demonstrated on the problem of output identification and control of a nonlinear two-mass torsional system.**

I. INTRODUCTION

Despite the large amount of literature on optimal control, there are not too much works regarding to optimization of initial conditions of dynamic systems. A reason is relatively limited applicability of such optimization methods. There are several applications in the area of chemical engineering, [1], [2], on the example of a fed-batch fermentation process where the problem is optimization of initial glucose concentration and initial volume such that the ethanol production rate is maximized. The second example is the problem of Earth-to-Mars optimal orbit transfer [3], where the minimum time transfer between orbits depends on the initial phase angle of the spacecraft with respect to Earth and on the initial phase angle of Mars with respect to Sun.

There are several applications of initial conditions optimization in linear control systems design. The problem of initial value compensation have appeared mostly in the disk drive community where servo systems need to satisfy very high performance demands, [4], [5], [6]. The problem is to find the optimal initialization of the controller integral state that gives the minimal achievable settling time without a pronounced overshoot. A similar problem is optimal initialization of the observer state, [7]. The observer initial conditions in industrial applications are usually set to zero, by default. When optimum initial conditions are used, rather than zero initial conditions, a significant improvement in the observer's transients may result.

On the other hand, the parameter optimization of dynamic systems has much wider applicability, especially for optimal tuning of controller gains, [8], [9]. In [10] parameters optimization of several different controllers for ship navigation is realized using genetic algorithms. In [11] parameters optimization of a sliding mode controller for submarine navigation is realized using simulated annealing and genetic algorithms. Also, the parameter identification problem can be formulated as parameter optimization problem, [12], [13].

The methods for parameter and initial conditions optimization can be classified in two main groups: a) deterministic nonlinear programming (NLP) based methodss, [14]; and b) stochastic/evolutionary based methods, [8]. The NLP-based methods are characterized by a large and sparse structure of Jacobian and Hessian matrices, so that they can be computationally expensive for large systems. On the other hand, the main drawback of stochastic/evolutionary based methods is slow convergence and long computational time, [1], [15], [16].

This paper presents a new numerical approach to initial conditions optimization which is based on the backpropagation-through-time (BPTT) algorithm [17]. The proposed algorithm is an extension of previous work where a similar approach is applied on open-loop optimal control [18] and robust feedback control [19]. The core of the proposed algorithm is exact gradient calculation of a terminal cost function with respect to optimized initial conditions. In contrast to NLP approach, Jacobian has no sparse structure and conjugate gradient modification of the algorithm provide much faster convergence in comparison with stochastic and evolutionary algorithms.

Further, it is shown that the parameter optimization problem can be easily transformed to initial conditions optimization problem. A benefit of such an indirect approach to parameter optimization is that final algorithm has simple structure which involve calculation of only one Jacobian. The second benefit is that the parameter optimization and initial conditions optimization can be easily combined. This is illustrated on the problem of output identification of nonlinear dynamic system with unknown nonzero initial conditions and on the problem of output dynamic controller design with free initial conditions.

This paper is organized as follows. The initial conditions optimization problem formulation and algorithm derivation are presented in Section II. The application of proposed algorithm to parameter optimization, identification and controller design is presented in Section III. The simulation results are presented in Section IV. Finally, the concluding remarks are emphasized in Section V.

This work was supported by the National Scientific Foundation of Republic of Croatia under Grant No. 120-1201842-3048.

Josip Kasac, Vladimir Milic, Branko Novakovic, Dubravko Majetic and Danko Brezak are with University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture, Zagreb HR-10000, Croatia. Emails: josip.kasac@fsb.hr, vladimir.milic@fsb.hr, branko.novakovic@fsb.hr, dubravko.majetic@fsb.hr, danko.brezak@fsb.hr .

978-1-4673-2529-5/12/$31.00 ©2012 IEEE

II. Initial Conditions Optimization

A. Continuous-Time Problem Formulation

The problem is to find a set of initial conditions $\hat{\mathbf{x}}(0)$ for the continuous nonlinear dynamic systems

$$\dot{\bar{\mathbf{x}}}(t) = \mathbf{f}_0(\bar{\mathbf{x}}(t), t), \qquad \bar{\mathbf{x}}(0) = \begin{bmatrix} \tilde{\mathbf{x}}(0) \\ \hat{\mathbf{x}}(0) \end{bmatrix}, \tag{1}$$

which minimize the cost function

$$J_0 = \Phi_0(\bar{\mathbf{x}}(t_f)) + \int_0^{t_f} F_0(\bar{\mathbf{x}}(t), t)dt, \tag{2}$$

subject to the initial and final conditions on the state vector

$$\mathbf{b}(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(t_f)) = \mathbf{0}, \tag{3}$$

and subject to the state vector inequality constraints

$$\mathbf{g}(\bar{\mathbf{x}}(t), t) \geq \mathbf{0}, \tag{4}$$

where $\bar{\mathbf{x}}(t) \in \mathbb{R}^{n_0}$ is the state vector, $\tilde{\mathbf{x}}(0) \in \mathbb{R}^{n_1}$ is the vector of fixed initial conditions, $\hat{\mathbf{x}}(0) \in \mathbb{R}^{n_2}$ is the vector of free initial conditions, where $n_0 = n_1 + n_2$, $\mathbf{b}(\cdot) \in \mathbb{R}^{n_b}$ is the vector function of boundary constraints, $\mathbf{g}(\cdot) \in \mathbb{R}^{n_g}$ is the vector function of parameter vector inequality constraints, $\Phi_0(\cdot) \in \mathbb{R}$ is the terminal cost function, $F_0(\cdot) \in \mathbb{R}$ is the subintegral function, and t_f is the terminal time.

B. Transformation to Terminal-Time Optimization Problem

The optimization problem (1)-(4) can be transformed to the problem of finding the set of initial conditions $\hat{\mathbf{x}}(0)$ for the systems

$$\begin{bmatrix} \dot{\bar{\mathbf{x}}}(t) \\ \dot{z}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{f}_0(\bar{\mathbf{x}}(t), t) \\ F(\bar{\mathbf{x}}(t), t) \end{bmatrix}, \quad \bar{\mathbf{x}}(0) = \begin{bmatrix} \tilde{\mathbf{x}}(0) \\ \hat{\mathbf{x}}(0) \end{bmatrix}, \quad z(0) = 0, \tag{5}$$

which minimize the terminal cost function

$$J = \Phi_0(\bar{\mathbf{x}}(t_f)) + z(t_f) + \sum_{k=1}^{n_b} K_{b,k} b_k^2(\bar{\mathbf{x}}(0), \bar{\mathbf{x}}(t_f)), \tag{6}$$

where $z(t_f)$ is the integral part of the cost function (2), and

$$F(\bar{\mathbf{x}}, t) = F_0(\bar{\mathbf{x}}, t) + \sum_{k=1}^{n_g} K_{g,k} g_k^2(\bar{\mathbf{x}}, t) H^-(g_k(\bar{\mathbf{x}}, t)), \tag{7}$$

and where $H^-(\xi)$ is the Heaviside step function defined as follows

$$H^-(\xi) = \begin{cases} 0, & \text{if} \quad \xi \geq 0 \\ 1, & \text{if} \quad \xi < 0 \end{cases} \tag{8}$$

while $K_{b,k}$ and $K_{g,k}$ are the coefficients of the penalty function for the boundary and inequality constraints, respectively.

Hence, the final continuous-time optimization problem is to find the set of initial conditions $\hat{\mathbf{x}}(0)$ for the systems

$$\dot{\mathbf{x}}(t) = \hat{\mathbf{f}}(\mathbf{x}(t), t), \qquad \mathbf{x}(0) = \begin{bmatrix} \tilde{\mathbf{x}}^T(0) & \hat{\mathbf{x}}^T(0) & 0 \end{bmatrix}^T, \tag{9}$$

which minimize the terminal cost function

$$J = \Phi(\mathbf{x}(0), \mathbf{x}(t_f)), \tag{10}$$

where the new state vector $\mathbf{x}(t) \in \mathbb{R}^n$, and new vector function $\hat{\mathbf{f}}(\cdot) \in \mathbb{R}^n$, $n = n_0 + 1$, are introduced

$$\mathbf{x}(t) = \begin{bmatrix} \bar{\mathbf{x}}(t) \\ z(t) \end{bmatrix}, \qquad \hat{\mathbf{f}}(\mathbf{x}(t), t) = \begin{bmatrix} \mathbf{f}_0(\bar{\mathbf{x}}(t), t) \\ F(\bar{\mathbf{x}}(t), t) \end{bmatrix}. \tag{11}$$

C. Discrete-Time Optimization Problem

The next step is time discretization of the optimization problem (9)-(10). The discrete time instances $t_i = i\tau$, $i = 0, 1, ..., N-1$, are introduced, where $\tau = t_f/N$ is the sampling interval, and N is the number of sampling intervals.

The final discrete-time optimization problem is to find the set of initial conditions $\hat{\mathbf{x}}(0)$ for the systems

$$\mathbf{x}(i+1) = \mathbf{f}(\mathbf{x}(i), i), \qquad \mathbf{x}(0) = \begin{bmatrix} \tilde{\mathbf{x}}^T(0) & \hat{\mathbf{x}}^T(0) & 0 \end{bmatrix}^T, \tag{12}$$

which minimize the terminal cost function

$$J = \Phi(\mathbf{x}(0), \mathbf{x}(N)), \tag{13}$$

for $i = 0, 1, ..., N-1$, where $\mathbf{x}(i) \equiv \mathbf{x}(t_i) = \mathbf{x}(i\tau)$, and

$$\mathbf{f}(\mathbf{x}(i), i) = \mathbf{x}(i) + \tau\hat{\mathbf{f}}(\mathbf{x}(i), i). \tag{14}$$

The accuracy of Euler integration method (12), (14) can be improved using the k-th order multistep Adams methods, which can be conveniently transformed into the equivalent discrete-time state-space form (12) with $\mathbf{f}(\cdot) \in \mathbb{R}^{nk}$, [20].

D. Gradient Calculation

The gradient descent algorithm according to the initial state vector is given as follows:

$$\mathbf{x}^{(l+1)}(0) = \mathbf{x}^{(l)}(0) - \eta^{(l)}\mathbf{\Omega}\frac{\partial J}{\partial \mathbf{x}^{(l)}(0)} \tag{15}$$

where index l represents the l-th iteration of the gradient algorithm, $l = 1, 2, \ldots, M$, $\eta^{(l)}$ is the learning rate, and M is the number of iterations of the gradient algorithm. The matrix $\mathbf{\Omega}$ has diagonal entries

$$\Omega_{jj} = \begin{cases} 0, & \text{if} \quad x_j(0) \text{ is fixed} \\ 1, & \text{if} \quad x_j(0) \text{ is free} \end{cases} \tag{16}$$

The gradient of the cost function (13) in the l-th iteration of the gradient algorithm and i-th sampling interval is given by

$$\frac{\partial J}{\partial x_j(0)} = \sum_{k=1}^{n} \frac{\partial J}{\partial x_k(N)} \frac{\partial x_k(N)}{\partial x_j(0)} \tag{17}$$

where $j = 1, 2, ..., n$. The partial derivatives $\frac{\partial x_k(N)}{\partial x_j(0)}$ can be calculated backward in time, starting from $N-1$:

$$\frac{\partial x_k(N)}{\partial x_j(N-1)} = \frac{\partial f_k(N-1)}{\partial x_j(N-1)} \tag{18}$$

where $f_k(i) \equiv f_k(\mathbf{x}(i))$, and $k = 1, 2, ..., n$. Further, for $i = N-2$:

$$\begin{aligned} \frac{\partial x_k(N)}{\partial x_j(N-2)} &= \sum_{p=1}^{n} \frac{\partial f_k(N-1)}{\partial x_p(N-1)} \frac{\partial x_p(N-1)}{\partial x_j(N-2)} \\ &= \sum_{p=1}^{n} \frac{\partial f_k(N-1)}{\partial x_p(N-1)} \frac{\partial f_p(N-2)}{\partial x_j(N-2)} \end{aligned} \tag{19}$$

and for $i = N-3$:

$$
\begin{aligned}
\frac{\partial x_k(N)}{\partial x_j(N-3)} &= \sum_{p=1}^{n} \frac{\partial f_k(N-1)}{\partial x_p(N-1)} \frac{\partial x_p(N-1)}{\partial x_j(N-3)} \\
&= \sum_{p=1}^{n} \frac{\partial f_k(N-1)}{\partial x_p(N-1)} \frac{\partial f_p(N-2)}{\partial x_j(N-3)} \\
&= \sum_{p=1}^{n} \frac{\partial f_k(N-1)}{\partial x_p(N-1)} \sum_{l=1}^{n} \frac{\partial f_p(N-2)}{\partial x_l(N-2)} \frac{\partial x_l(N-2)}{\partial x_j(N-3)} \\
&= \sum_{p=1}^{n} \frac{\partial f_k(N-1)}{\partial x_p(N-1)} \sum_{l=1}^{n} \frac{\partial f_p(N-2)}{\partial x_l(N-2)} \frac{\partial f_l(N-3)}{\partial x_j(N-3)}.
\end{aligned} \tag{20}
$$

By introducing matrices $\mathbf{X}(0), \mathbf{Z}(0) \in \mathbb{R}^{n\times n}$,

$$\mathbf{X}(i) = \frac{\partial \mathbf{f}(i)}{\partial \mathbf{x}(i)}, \quad \mathbf{Z}(i) = \frac{\partial \mathbf{x}(N)}{\partial \mathbf{x}(i)}, \tag{21}$$

the above derivatives $\partial x_k(N)/\partial x_j(i)$, $i = N-1, N-2, N-3$, can be expressed in a more compact matrix form as

$$\mathbf{Z}(N-1) = \mathbf{X}(N-1), \tag{22}$$

$$\mathbf{Z}(N-2) = \mathbf{X}(N-1)\cdot\mathbf{X}(N-2), \tag{23}$$

$$\mathbf{Z}(N-3) = \mathbf{X}(N-1)\cdot\mathbf{X}(N-2)\cdot\mathbf{X}(N-3). \tag{24}$$

This procedure can be further continued as follows

$$\mathbf{Z}(N-i) = \prod_{k=0}^{i-1} \mathbf{X}(N-1-k), \tag{25}$$

for $i = 1, 2, ..., N$.

By introducing matrices

$$\mathbf{J}_x(0) = \frac{\partial J}{\partial \mathbf{x}(0)}, \quad \mathbf{J}_x(N) = \frac{\partial J}{\partial \mathbf{x}(N)}, \tag{26}$$

the final gradient $\mathbf{J}_x(0)$ can be computed by the following backward-in-time recursive matrix relations:

$$\mathbf{J}_x(0) = \mathbf{Z}^T(0)\cdot\mathbf{J}_x(N), \tag{27}$$

$$\mathbf{Z}(i) = \mathbf{D}(i)\cdot\mathbf{X}(i), \tag{28}$$

$$\mathbf{D}(i) = \mathbf{D}(i+1)\cdot\mathbf{X}(i+1), \tag{29}$$

for $i = N-2,\ N-3,\ ...,\ 0$, with the initial condition

$$\mathbf{D}(N-1) = \mathbf{I}, \tag{30}$$

where $\mathbf{I} \in \mathbb{R}^{n\times n}$ is the unit matrix.

The convergence properties of the standard gradient algorithm (15) can be significantly improved using conjugate gradient methods [18], [21].

III. Parameter Optimization Problem

The algorithm for optimization of initial conditions can be applied to different optimization problems, like parameter optimization which further can be used for parameter identification and output controller design. In this section we will demonstrate how to transform different optimization problems to initial conditions optimization problem formulation (1)-(4).

A. Transformation to Parameter Optimization Problem

The problem is to find a set of parameters $\mathbf{p} \in \mathbb{R}^{n_p}$ for the continuous nonlinear dynamical systems

$$\dot{\tilde{\mathbf{x}}}(t) = \tilde{\mathbf{f}}(\tilde{\mathbf{x}}(t), \mathbf{p}, t), \quad \tilde{\mathbf{x}}_0 = \tilde{\mathbf{x}}(0) \tag{31}$$

which minimize the cost function

$$J = \Phi_0(\tilde{\mathbf{x}}(t_f), \mathbf{p}) + \int_0^{t_f} F_0(\tilde{\mathbf{x}}(t), \mathbf{p}, t)dt, \tag{32}$$

subject to the final conditions on the state vector

$$\mathbf{b}(\tilde{\mathbf{x}}(t_f), \mathbf{p}) = \mathbf{0}, \tag{33}$$

and subject to the state vector inequality constraints

$$\mathbf{g}(\tilde{\mathbf{x}}(t), \mathbf{p}, t) \geq \mathbf{0}, \tag{34}$$

where $\tilde{\mathbf{x}}(t) \in \mathbb{R}^{n_0}$ is the state vector, $\mathbf{b}(\cdot) \in \mathbb{R}^{n_b}$ is the vector function of boundary constraints, $\mathbf{g}(\cdot) \in \mathbb{R}^{n_g}$ is the vector function of inequality constraints, and t_f is the terminal time.

If we introduce additional state variable $\hat{\mathbf{x}}(t) = \mathbf{p}$, then system (31) can be rewritten as

$$\dot{\tilde{\mathbf{x}}}(t) = \tilde{\mathbf{f}}(\tilde{\mathbf{x}}(t), \hat{\mathbf{x}}(t), t), \quad \tilde{\mathbf{x}}_0 = \tilde{\mathbf{x}}(0), \tag{35}$$

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{0}, \quad \hat{\mathbf{x}}_0 = \hat{\mathbf{x}}(0), \tag{36}$$

and parameter optimization problem is equivalent to optimization of set of initial conditions $\hat{\mathbf{x}}(0)$. By introducing

$$\bar{\mathbf{x}}(t) = \begin{bmatrix} \tilde{\mathbf{x}}(t) \\ \hat{\mathbf{x}}(t) \end{bmatrix}, \quad \mathbf{f}_0(\bar{\mathbf{x}}(t), t) = \begin{bmatrix} \tilde{\mathbf{f}}(\bar{\mathbf{x}}(t), t) \\ \mathbf{0} \end{bmatrix}, \tag{37}$$

the parameter optimization problem (31)-(34) formally became equivalent to initial conditions optimization problem (1)-(4).

B. Parameter Identification Problem Formulation

The parameter identification problem can be considered as a special case of the parameter optimization problem (31)-(32). The problem is to find the parameter vector $\mathbf{p}$ of the dynamic system

$$\dot{\tilde{\mathbf{x}}}(t) = \tilde{\mathbf{f}}(\tilde{\mathbf{x}}(t), \mathbf{u}(t), \mathbf{p}), \quad \tilde{\mathbf{x}}(0) = \tilde{\mathbf{x}}_0, \tag{38}$$

$$\mathbf{y}(t) = \mathbf{h}_0(\tilde{\mathbf{x}}(t)), \tag{39}$$

which will minimize the cost function

$$J = \int_0^{t_f} (\mathbf{y}(t) - \mathbf{y}_m(t))^T \mathbf{Q} (\mathbf{y}(t) - \mathbf{y}_m(t))dt, \tag{40}$$

where $\mathbf{u}(t) \in \mathbb{R}^m$ is known input excitation vector function, $\mathbf{y}_m(t) \in \mathbb{R}^p$ is measured output vector over interval $[0, t_f]$, and $\mathbf{Q} \in \mathbb{R}^{p\times p}$ is the weighting matrix.

The input vector function $\mathbf{u}(t)$ should be persistent excitation signal which provide excitation of all system modes. Such a signal is pseudo-random signal or normally distributed random signal.

C. Output Controller Design Problem Formulation

The optimal output controller design can be formulated on the following way. For the nonlinear control system

$$\dot{\breve{\mathbf{x}}}(t) = \breve{\mathbf{f}}(\breve{\mathbf{x}}(t), \mathbf{u}(t), t), \qquad \breve{\mathbf{x}}(0) = \breve{\mathbf{x}}_0, \tag{41}$$
$$\mathbf{y}(t) = \mathbf{h}_0(\breve{\mathbf{x}}(t)), \tag{42}$$

find the dynamic controller

$$\mathbf{u}(t) = \mathbf{h}_1(\mathbf{y}(t), \mathbf{z}(t), \mathbf{k}), \tag{43}$$
$$\dot{\mathbf{z}}(t) = \mathbf{h}_2(\mathbf{y}(t), \mathbf{z}(t), \mathbf{k}), \qquad \mathbf{z}(0) = \mathbf{z}_0, \tag{44}$$

which will minimize cost function

$$J = \int_0^{t_f} F_0(\breve{\mathbf{x}}(t), \mathbf{u}(t), t)dt. \tag{45}$$

The dynamic controller is represented by some vector functions $\mathbf{h}_1(\cdot) \in \mathbb{R}^{n_m}$, $\mathbf{h}_2(\cdot) \in \mathbb{R}^{n_z}$, which are parameterized by the vector of controller gains $\mathbf{k} \in \mathbb{R}^{n_k}$.

The control system (41)-(44) can be rewritten as

$$\dot{\breve{\mathbf{x}}} = \breve{\mathbf{f}}\left(\breve{\mathbf{x}}, \mathbf{h}_1(\mathbf{h}_0(\breve{\mathbf{x}}), \mathbf{z}, \mathbf{k}), t\right), \quad \breve{\mathbf{x}}(0) = \breve{\mathbf{x}}_0, \tag{46}$$
$$\dot{\mathbf{z}} = \mathbf{h}_2(\mathbf{h}_0(\breve{\mathbf{x}}), \mathbf{z}, \mathbf{k}), \qquad \mathbf{z}(0) = \mathbf{z}_0, \tag{47}$$

and cost function (45) as

$$J = \int_0^{t_f} F_0(\breve{\mathbf{x}}(t), \mathbf{h}_1(\mathbf{h}_0(\breve{\mathbf{x}}(t)), \mathbf{z}(t), \mathbf{k}), t)dt. \tag{48}$$

The output design problem (46)-(48) is a special case of the parameter optimization problem (31)-(32), where $\mathbf{p} = \mathbf{k}$ and

$$\tilde{\mathbf{x}}(t) = \begin{bmatrix} \breve{\mathbf{x}}(t) \\ \hat{\mathbf{z}}(t) \end{bmatrix}, \qquad \tilde{\mathbf{f}}(\tilde{\mathbf{x}}(t), \mathbf{p}, t) = \begin{bmatrix} \breve{\mathbf{f}}(\tilde{\mathbf{x}}(t), \mathbf{p}, t) \\ \mathbf{h}_2(\tilde{\mathbf{x}}(t), \mathbf{p}) \end{bmatrix}. \tag{49}$$

IV. Simulation Example

In this section we will demonstrate the application of proposed algorithm to the problem of identification and output control of a two-mass torsional nonlinear mechanical system. The two masses are weakly coupled by nonlinearly parameterized torsional spring torques, making the problem challenging both for identification and output control.

The Jacobian $\mathbf{X}(i)$ is calculated numerically using second-order difference scheme [20], and a modified Fletcher-Reeves conjugate gradient algorithm is used [21].

A. Dynamic Model of Two-mass Torsional System

The nonlinear dynamic model of two-mass torsional system is

$$\dot{x}_1 = x_2, \tag{50}$$
$$\dot{x}_2 = \phi(x_1; p_1, p_2) - p_3 x_2 + \phi(x_1 - x_3; p_7, p_8), \tag{51}$$
$$\dot{x}_3 = x_4, \tag{52}$$
$$\dot{x}_4 = \phi(x_3; p_4, p_5) - p_6 x_4 - \phi(x_1 - x_3; p_7, p_8) + p_9 u, \tag{53}$$

where $y_1 = x_1$ and $y_2 = x_3$ are measured output positions, while x_2 and x_4 are unmeasurable velocities of the first and second mass, respectively. The external torque u is acting directly only on the second mass. The function

$$\phi(\xi; \rho_1, \rho_2) = -\rho_1 \xi e^{-\rho_2 \xi}, \tag{54}$$

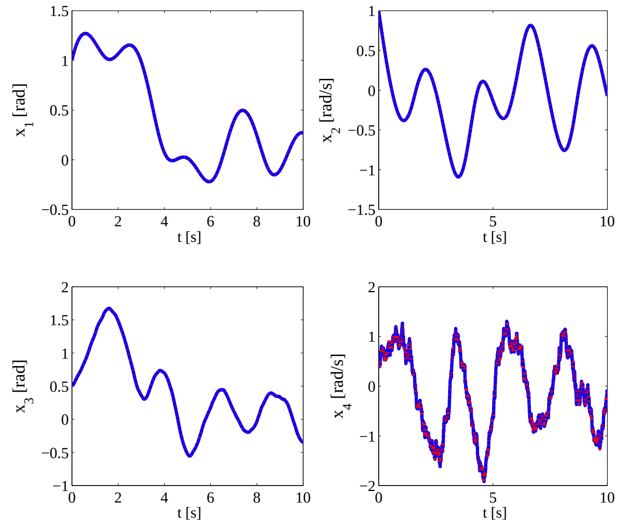

Fig. 1. The system responses for true and estimated parameters.

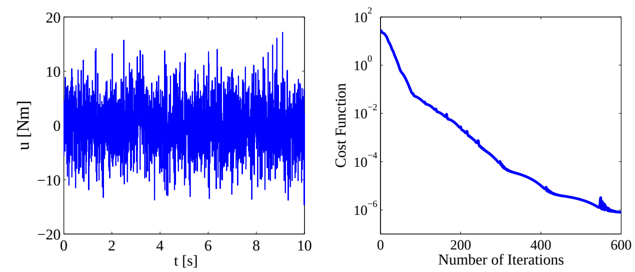

Fig. 2. The input excitation variable (left) and cost function depending on the number of iterations of gradient algorithm (right).

represents the nonlinear torsional spring torque parameterized linearly by parameter ρ_1 and nonlinearly by parameter ρ_2, [22]. The overall nonlinear state-space model is parameterized by nine parameters p_i, $i = 1, 2, ..., 9$, which values are shown in Table I.

B. Parameter Identification

The first problem is parameter identification of the dynamic model (50)-(53) based on the output measurement of system positions only $\mathbf{y}_m(t) = [x_{1m}\ x_{3m}]^T$, under external excitation $u(t)$. Additional assumption is that system initial conditions are not equal to zero. Since the system velocities are not measurable, that means that unknown initial velocities also should be estimated.

The terminal time is $t_f = 10\,\mathrm{s}$ and the sampling interval is $\tau = 0.005\,\mathrm{s}$, so that number of time intervals is $N = 2000$. The number of iterations of the Fletcher-Reeves conjugate gradient algorithm is $M = 600$. Fig. 1. shows the response of the two mass system (50)-(53) in the case of true and estimated values of parameters, which are listed in Table I. For the input excitation variable $u(t)$ is chosen normally distributed random signal, shown in Fig. 2., since it provides fast convergence of estimated parameters toward true values

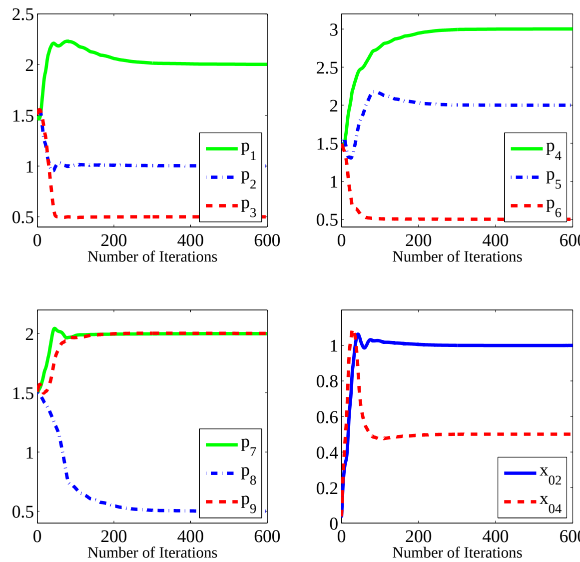

Fig. 3. Estimated parameters and initial conditions depending on the number of iterations of gradient algorithm.

TABLE I
TRUE AND ESTIMATED VALUES OF SYSTEM PARAMETERS AND INITIAL CONDITIONS.

Parameter notation	True value	Estimated value	Relative error (%)
p_1	2.0	2.0020	0.1015
p_2	1.0	1.0011	0.1143
p_3	0.5	0.4999	0.0252
p_4	3.0	3.0009	0.0295
p_5	2.0	1.9995	0.0232
p_6	0.5	0.5004	0.0713
p_7	2.0	1.9993	0.0340
p_8	0.5	0.5004	0.0756
p_9	2.0	2.0012	0.0593
x_{02}	1.0	0.9997	0.0342
x_{04}	0.5	0.5003	0.0588

as shown in Fig. 3. From the Table I we can see that relative error of estimated parameters and initial values does not exceed 0.1%.

C. Output Controller Design

We consider the problem of set point control of two-mass system where the goal is to stabilize the position of the first mass x_1 in desired reference state x_{1d}, under assumption that only output $\mathbf{y}(t) = [x_1 \; x_3]^T$ is measurable. The output controller in a form of a saturated PID controller with first-

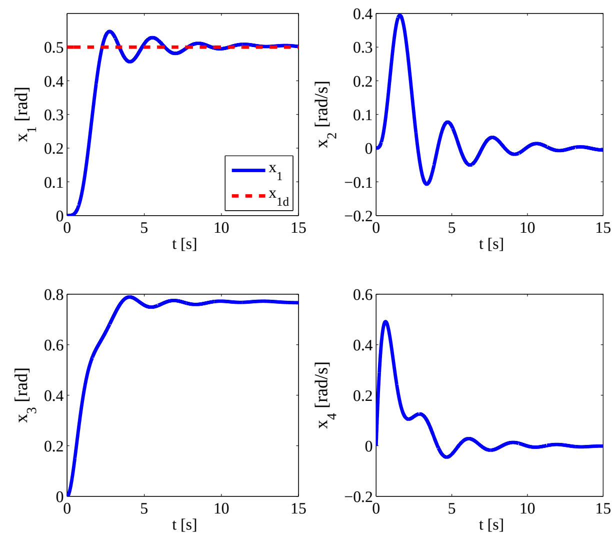

Fig. 4. The system responses for optimal values of controller gains.

order velocity estimation filter is proposed

$$u(t) = \tanh(-k_1(x_1 - x_{1d}) + k_2 z_2 - k_3 z_1), \quad (55)$$
$$\dot{z}_1 = x_1 - x_{1d}, \quad (56)$$
$$\dot{z}_2 = -k_4 z_2 + k_5 x_3, \quad (57)$$

where the parameters k_i, $i = 1, 2, ..., 5$, are controller gains and $z_2(0)$ is free initial condition of the filter which should be determined by minimizing the cost function (45) in the form

$$J = \int_0^{t_f} e^2(q_1 + q_2 \mathrm{sign}(x_{1d} e))dt. \quad (58)$$

where $e = x_1 - x_{1d}$, and $q_1, q_2, q_1 - q_2 > 0$. The cost function (58) is an asymmetric integral square error (ISE) performance index which additionally penalizes the overshot of output position x_1 by rate $\mu = (q_1 + q_2)/(q_1 - q_2)$. In the case $q_2 = 0$, the cost function (58) becomes the standard ISE performance index.

The terminal time is $t_f = 15\,\mathrm{s}$ and the number of time intervals is $N = 4000$, so that sampling interval is $\tau = 0.0037\,\mathrm{s}$. The number of iterations of the Fletcher-Reeves conjugate gradient algorithm is $M = 600$. The cost function (58) parameters are chosen as $q_1 = 3$ and $q_2 = 2$, so that overshot penalization rate is $\mu = 5$. Fig. 4. shows the response of the two mass system (50)-(53) in the case of optimal values of controller gains, which are listed in Table II. The control variable $u(t)$ is shown in Fig. 5. Fig. 6 illustrates the convergence of the controller gains depending on the number of iterations of the conjugate gradient algorithm.

V. CONCLUSIONS

In this paper a gradient-based numerical approach for initial condition optimization of nonlinear dynamic systems is presented. The applicability of the proposed algorithm

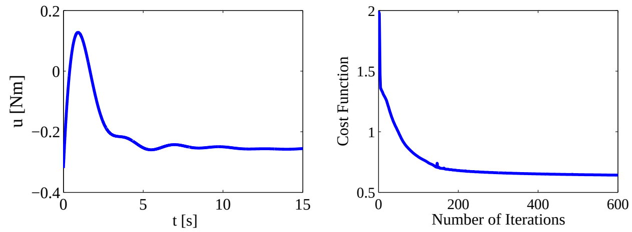

Fig. 5. The control variable (left) and cost function depending on the number of iterations of gradient algorithm (right).

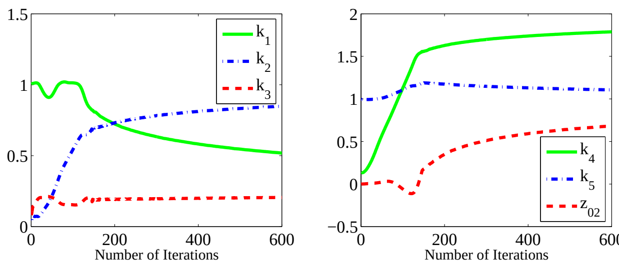

Fig. 6. Controller gains and initial condition depending on the number of iterations of gradient algorithm.

is extended to parameter optimization problem and demonstrated on examples of output identification and control of a two-mass torsional system. The optimization results have illustrated favorable features of the algorithm in terms of accuracy (e.g., a few thousands of time grid points can be used), consistent numerical stability, and relatively fast convergence properties.

The future research will be oriented towards extension of possible applications of proposed algorithm to decentralized robust controller design, minimum time feedback control and learning algorithms for dynamic neural networks. Also, the presented algorithm will be integrated with previously developed open-loop optimal control algorithm providing a unified framework for simultaneous optimization of control variables, parameters and initial conditions of nonlinear multivariable dynamic systems.

TABLE II

INITIAL AND OPTIMAL VALUES OF CONTROLLER PARAMETERS AND INITIAL CONDITIONS.

Parameter notation	Initial value	Optimal value
k_1	1.0	0.5186
k_2	0.1	0.8485
k_3	0.1	0.2056
k_4	0.1	1.7881
k_5	1.0	1.1059
z_{02}	0.0	0.6826

REFERENCES

[1] J.-P. Chiou and F.-S. Wang, "Hybrid method of evolutionary algorithms for static and dynamic optimization problems with application to a fed-batch fermentation process," *Computers and Chemical Engineering*, vol. 23, pp. 1277–1291, 1999.

[2] M. D. Kapadi and R. D. Gudi, "Optimal control of fed-batch fermentation involving multiple feeds using differential evolution," *Process Biochemistry*, vol. 39, pp. 1709–1721, 2004.

[3] A. Miele and T. Wang, "Optimal trajectories for earth-to-mars flight," *Journal of Optimization Theory and Applications*, vol. 95, pp. 467–499, 1997.

[4] T.-Y. Doh and J. R. Ryoo, "A linear matrix inequality approach to initial value compensation for mode switching control in hard disk drive servo systems," in *Proceedings of International Conference on Control, Automation and Systems*, Oct. 2008, pp. 1692 –1695.

[5] M. Johansson, "Optimal initial value compensation for fast settling times in mode-switching control systems," in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 5, 2000, pp. 5137–5142.

[6] M. Teixeira, N. Silva, E. Assuncao, and E. Machado, "Design of fuzzy regulators with optimal initial conditions compensation," in *Proc. of IEEE International Conf. on Fuzzy Systems*, 2006, pp. 84 –91.

[7] C. D. Johnson, "Optimal initial conditions for full-order observers," *International Journal of Control*, vol. 48, no. 3, pp. 857–864, 1988.

[8] A. Varsek, T. Urbancic, and B. Filipic, "Genetic algorithms in controller design and tuning," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 5, pp. 1330 –1339, 1993.

[9] C. Pedret, R. Vilanova, R. Moreno, and I. Serra, "A refinement procedure for PID controller tuning," *Computers and Chemical Engineering*, vol. 26, no. 6, pp. 903 – 908, 2002.

[10] E. Alfaro-Cid, E. McGookin, and D. Murray-Smith, "A comparative study of genetic operators for controller parameter optimisation," *Control Engineering Practice*, vol. 17, no. 1, pp. 185 – 197, 2009.

[11] E. W. McGookin and D. J. Murray-Smith, "Submarine manoeuvring controllers optimisation using simulated annealing and genetic algorithms," *Control Engineering Practice*, vol. 14, pp. 1 – 15, 2006.

[12] C. Medler and C.-C. Hsu, "An algorithm for nonlinear system parameter identification," *IEEE Transactions on Automatic Control*, vol. 14, no. 6, pp. 726 – 728, 1969.

[13] A. Mehta and H. Kaufman, "Sensitivity analysis for parameter identification using optimal control indices," in *Proceedings of IEEE International Conference on Control Applications*, Dearborn, MI, September 15-18 1996, pp. 564 –569.

[14] J. R. Banga, E. Balsa-Canto, C. G. Moles, and A. A. Alonso, "Dynamic optimization of bioprocesses: Efficient and robust numerical strategies," *Journal of Biotechnology*, vol. 117, p. 407 419, 2005.

[15] J. Barrera and J. Flores, "Search of initial conditions for dynamic systems using intelligent optimization methods," in *Proceedings of Electronics, Robotics and Automotive Mechanics Conference*, Los Alamitos, CA, USA, Sept. 2007, pp. 645 –650.

[16] T. K. Mok, H. Liu, Y. Ni, F. F. Wu, and R. Hui, "Tuning the fuzzy damping controller for upfc through genetic algorithm with comparison to the gradient descent training," *Electrical Power and Energy Systems*, vol. 27, pp. 275–283, 2005.

[17] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proceedings of IEEE*, vol. 78, pp. 1550–1560, 1990.

[18] J. Kasac, J. Deur, B. Novakovic, I. Kolmanovsky, and F. Assadian, "A conjugate gradient-based BPTT-like optimal control algorithm with vehicle dynamics control application," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 6, pp. 1587–1595, 2011.

[19] V. Milic, J. Kasac, D. Majetic, and Z. Situm, "A BPTT-like Min-Max optimal control algorithm for nonlinear systems," in *Proceedings of International Conference of Numerical Analysis and Applied Mathematics*, Rhodes, Greece, 19-25 September 2010.

[20] J. Kasac, J. Deur, B. Novakovic, and I. Kolmanovsky, "A BPTT-like optimal control algorithm with vehicle dynamics control application," in *Proceedings of ASME International Mechanical Engineering Congress and Exposition*, Boston, MA, 31 Sept. - 6 Nov. 2008.

[21] ——, "A conjugate gradient-based BPTT-like optimal control algorithm," in *Proceedings of IEEE Multi-conference on Systems and Control*, Saint Petersburg, Russia, 8-10 July 2009.

[22] L. Magni, G. De Nicolao, R. Scattolini, and F. Allgwer, "Robust model predictive control for nonlinear discrete-time systems," *International Journal of Robust and Nonlinear Control*, vol. 13, no. 3-4, pp. 229–246, 2003.