Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Implicit dual adaptive control for systems with functional uncertainties

Ladislav Král and Miroslav Šimandl and Miroslav Flídr

European Centre of Excellence "NTIS - New Technologies for the Information Society", Faculty of Applied Sciences, University of West Bohemia, Pilsen, Czech Republic, ladkral, simandl, flidr@ntis.zcu.cz

Abstract: The paper proposes an implicit type of dual control for a class of nonlinear stochastic systems subject to functional uncertainty. The unknown functions of the system are modelled by multi-layered perceptron neural networks where the unknown parameters are found in real-time. The control design is based on the Bellman optimisation recursion where the length of the recursion is shortened to two stages to reduce computational burdens and to ensure dual features between estimation and control aims. The inherent obstacle of determining the expectation is tackled by employing a technique based on the stochastic integration rule. The design is then accomplished using an iterative procedure, which is summarised by algorithms. Numerical simulations and a Monte Carlo analysis show that the proposed approach may compete with existing solutions based solely on the explicit type of dual control and removes their drawback of tuning additional design parameters.

1. INTRODUCTION

Adaptive control of nonlinear stochastic systems with unknown functions offers an interesting challenge [Fabri and Kadirkamanathan, 2001, Herzallah and Lowe, 2002, Sarangapani, 2006]. It can be understood as a natural effort of extension of adaptive control from the class of linear systems and nonlinear systems with unknown parameters to the complex systems with functional uncertainty. Sometimes this attractive direction of adaptive control is also called a functional adaptive control [Fabri and Kadirkamanathan, 2001].

In principle, an optimal solution to the functional adaptive control is known and leads to the use of dynamic programming to solve the Bellman optimisation recursion (BOR). Unfortunately, a derivation of analytical solutions is unavailable and a numerical solution is difficult to achieve because of the curse of dimensionality. Moreover, an application in adaptive control is often complicated by the need to address a tedious problem of nonlinear estimation, i.e. the necessity to solve the Bayesian recursive relations (BRR) to find a probabilistic description of the uncertainties in the model. Since it is impossible to find the optimal solution, much attention has been concentrated on various approximate approaches [Fabri and Kadirkamanathan, 2001, Filatov and Unbehauen, 2004]. One of the main directions is based on the stochastic control principles originated from the work of Fel'dbaum [1965], which pointed out a contrast between estimation and control aims. It means that the controller besides respecting the uncertainty in model knowledge to achieve tracking performance should excite the system to reduce the uncertainty in the future. Control design which respects these conflicting features, called caution and probing, is referred to as dual control [Fel'dbaum, 1965, Bar-Shalom and Tse, 1976]. An exhaustive classification of sub-optimal dual controllers can be found in [Filatov and Unbehauen, 2004]. The main idea of these methods lies in the design of adaptive systems that are not optimal but at least have the main dual features of optimal (adaptive) control systems.

Fundamentally, dual control approximations may be separated into two distinct categories: *explicit* and *implicit* ones. The *explicit* type dual control methods can be seen as simpler and straightforward, which allows finding a closed-form solution. They are typically based on a minimisation of one-step ahead cost functions accompanied by including a term that induces the probing signal so that the system maintains the dual properties [Wieselander and Wittenmark, 1971, Milito et al., 1982, Filatov and Unbehauen, 2004]. The *implicit* type dual controllers contain all procedures which employ various approximations of either the Bellman function or the probability density functions [Tse et al., 1973, Bayard and Schumitzky, 2010]. The main characteristic of these controllers is that they provide a good performance, but are more complicated for practical implementation and often have higher computational demands.

Most existing methods of dual control are designed mainly for linear models in both state and input-output representation. The functional approach has been so far focused especially on algorithms having a practical on-line implementation. The original concept was designed in Fabri and Kadirkamanathan [1998], where the innovations dual control criterion, originally proposed by Milito et al. [1982], was used as a cost function together with two types of neural networks as a model for the unknown nonlinear functions. This was followed by Šimandl et al. [2005] which utilised Bicriterial cost function [Filatov and Unbehauen, 2004] instead of the IDC, and a Gaussian Sum Filter for parameter estimation. In Bugeja et al. [2009], it was tested in a practical control task of the nonholonomic wheeled mobile robot, and the functional approach was successfully extended to a more general MIMO class of nonlinear systems by Král and Šimandl [2011b] and Fabri and Bugeja [2013]. Further efforts tackle the individual problems such as the design procedure with predictive control extension [Král and Šimandl, 2011a], an improvement of estimation accuracy [Fabri and

⋆ The work was supported by the European Regional Development Fund (ERDF), project "NTIS - New Technologies for the Information Society", European Centre of Excellence, CZ.1.05/1.1.00/02.0090 and by the Czech Science Foundation, project GACR P103/13/07058J.

Bugeja, 2013] or relaxed a minimum phase property [Král and Šimandl, 2013].

A common feature of all above mentioned solutions to the functional approach is that they are based on the *explicit* dual control methods only. The common drawback of these solutions can be treated by (i) the algorithms include several tuning parameters that are selected ad-hoc or by trial-and-error [Filatov and Unbehauen, 2004], (ii) derivation of the control law depends on a type of estimator of unknown parameters [Fabri and Bugeja, 2013] and (iii) they are based on rough approximations of the original optimal problem and so the control performance remains, therefore, inadequate. So far, the implicit approach is a neglected group of methods, although it has been shown that it can bring a significant improvement in control quality. They are ignored mainly because they are computationally demanding, applicable only to systems with low rate uncertainty [Lee and Lee, 2009] or require the implementation of an intensive off-line identification process [Herzallah and Lowe, 2008]. To the best knowledge of the authors, there is no solution to the functional approach based on implicit methods of dual control.

Based on the above motivation, the general goal of the paper is to design an implicit dual control (IMDUC) for stochastic systems with functional uncertainties. The design shall follow the previous co-authors paper [Flídr and Šimandl, 2013], where a new IMDUC was proposed for a linear stochastic system in state space representation. This concept tackles solvability of the BOR by a reformulation of the problem to the receding horizon type of the optimisation, and by reduction of the BOR to two stage where the inherent expectations were determined by employing the stochastic integration rule. The main novel contribution of the presented paper is the extension of the IMDUC developed by Flídr and Šimandl [2013] to a generic class of nonlinear systems with unknown functions. The proposed controller apparently represents the first attempt to implement the implicit type of dual controller for such a complex system.

The rest of the paper will present a description of the problem formulation in Section 2, followed by a control design based on the solution to the optimisation recursion and the stochastic integration rule in Section 3. A numerical example demonstrates a control quality is contained in Section 4 and finally, Section 5 concludes the paper.

2. PROBLEM FORMULATION

The dynamical system to be controlled is a nonlinear stochastic discrete time-invariant system given in an input-output representation as

$$\mathcal{S}: \qquad y_{k+1} = f(\boldsymbol{x}_k) + g(\boldsymbol{x}_k)u_k + e_{k+1}, \tag{1}$$

where $f, g : \mathbb{R}^{n_y+n_u} \to \mathbb{R}$ are unknown nonlinear functions, $\boldsymbol{x}_k = [y_{k-n_y+1}, \ldots, y_k, u_{k-n_u}, \ldots, u_{k-1}]^T \in \mathbb{R}^{n_y+n_u}$ is the state vector, u_k and y_k are input and output signals at discrete time instants $k \in 0, 1, \ldots,$ N-1 and $\{e_k\}$ is an additive noise and the following assumptions are considered:

Assumption 1: The nonlinear functions $f(\boldsymbol{x}_k), g(\boldsymbol{x}_k) \in C^\infty$.

Assumption 2: The structural parameters n_y and n_u of the system are known.

Assumption 3: The system has a globally uniformly asymptotically stable zero dynamics and the nonlinear function $g(\boldsymbol{x}_k)$ is bounded away from zero for all $\boldsymbol{x}_k$ [Chen and Khalil, 1995].

Assumption 4: $\{e_k\} \in \mathbb{R}$ is a Gaussian sequence with a known mean μ and variance σ^2.

Since the nonlinear functions $f(\boldsymbol{x}_k)$, $g(\boldsymbol{x}_k)$ are unknown, an appropriate model of the system (1) together with a searching process for the optimal parameter values has to be found. Models based on Multi-Layer Perceptron (MLP) neural networks (NN) are preferred as a suitable compromise between complexity and accuracy of the model. Furthermore the using NN is justified by the validity of universal approximation property [Haykin, 1999]. Then, a generalised parametric model structure of the system (1) can be used in the following state space form

$$\mathcal{M}: \qquad \boldsymbol{\Theta}_{k+1} = \boldsymbol{\Theta}_k \tag{2}$$

$$y_{k+1} = h(\boldsymbol{\Theta}_k, \boldsymbol{x}_k, u_k) + e_{k+1}, \tag{3}$$

where $\boldsymbol{\Theta}_k \in \mathbb{R}^{n_\Theta}$ is a vector of the unknown parameters of the NN based model, which are assumed to be t-invariant. Further, the nonlinear function $h : \mathbb{R}^{n_y+n_u+1} \to \mathbb{R}$ is defined as

$$h(\boldsymbol{\Theta}_k, \boldsymbol{x}_k, u_k) = \hat{f}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k) + \hat{g}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k)u_k, \tag{4}$$

where $\hat{f}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k)$ and $\hat{g}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k)$ represent a pair of the NN which approximate the unknown functions of the system. They can be described as

$$\hat{f}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k) = \left(\boldsymbol{c}_k^f\right)^T \left[\left(\phi\left(\boldsymbol{x}_k, \boldsymbol{W}_k^f\right)\right)^T, 1\right]^T, \tag{5}$$

$$\hat{g}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k) = \left(\boldsymbol{c}_k^g\right)^T \left[\left(\phi\left(\boldsymbol{x}_k, \boldsymbol{W}_k^g\right)\right)^T, 1\right]^T, \tag{6}$$

where ϕ is a chosen activation function of the NN and $\boldsymbol{W}_k^f = \left[(\boldsymbol{w}_{1,k}^f)^T, \ldots, (\boldsymbol{w}_{nf,k}^f)^T \right]$, $\boldsymbol{W}_k^g = \left[(\boldsymbol{w}_{1,k}^g)^T, \ldots, (\boldsymbol{w}_{ng,k}^g)^T \right]$ are the hidden layers weights and $\boldsymbol{c}_k^f$ and $\boldsymbol{c}_k^g$ are the output layers weights of the NN $\hat{f}$ and $\hat{g}$, with the vector $\boldsymbol{w}_{i,k}^j$ denoting weights between inputs and the i-th neuron in a hidden layer of the j-th neural network.

Let all the unknown parameters of the NN be included into a single parameter vector $\boldsymbol{\Theta}_k$ in the compact form as

$$\boldsymbol{\Theta}_k = \left[(\boldsymbol{c}_k^f)^T, (\boldsymbol{w}_{1,k}^f)^T, \ldots (\boldsymbol{w}_{nf,k}^f)^T, (\boldsymbol{c}_k^g)^T, (\boldsymbol{w}_{1,k}^g)^T, \ldots (\boldsymbol{w}_{ng,k}^g)^T \right]^T \tag{7}$$

with its length denoted by n_Θ. Moreover, it is considered that $\boldsymbol{\Theta}_k$ is modelled as a random variable with Gaussian distribution

$$p\left(\boldsymbol{\Theta}_k | \boldsymbol{I}^k\right) \approx \mathcal{N}\left\{\boldsymbol{\Theta}_k : \hat{\boldsymbol{\Theta}}_k, \boldsymbol{P}_k\right\}, \tag{8}$$

with known initial conditions given by the mean $\hat{\boldsymbol{\Theta}}_0$ and the covariance matrix $\boldsymbol{P}_0$. The symbol $\boldsymbol{I}^k$ represents complete information available up to the time instant k , i.e.

$$\boldsymbol{I}^k = (\boldsymbol{u}_0^{k-1}, \boldsymbol{y}_0^k), \tag{9}$$

where $\boldsymbol{u}_0^k = [u_0, \ldots, u_k]$ and $\boldsymbol{y}_0^k = [y_0, \ldots, y_k]$.

Equations (2)–(7) describe a nonlinear stochastic state space model of the system (1). Unfortunately, dependence of y_{k+1} on the parameters of model $\boldsymbol{\Theta}_k$ is nonlinear. Therefore, it is advisable to exploit a convenient method to compute the probability distribution of these variables and its propagation through time. Although it is possible to use a variety of the Kalman filtering methods [Haykin, 2001], the extended Kalman Filter (EKF) is considered as a nonlinear estimator for its practicality and computationally moderateness.

Since the parameter vector $\boldsymbol{\Theta}_k$ is unknown, the problem becomes one of the adaptive control problems. The aim of the optimal adaptive control problem is to find the control law

$$u_k = \underset{u_k}{\operatorname{argmin}} J_k(\boldsymbol{I}^k) \qquad k = 0, 1, \ldots, N-1 \tag{10}$$

with the criterion J_k specified as

$$J_k = E\left\{ \sum_{i=k}^{N-1} \mathcal{L}_i(y_{i+1}, u_i, \boldsymbol{\Theta}_i) \right\}, \tag{11}$$

where $\mathcal{L}_i$ is a cost functional, the conditional expectation operator E is taken over all underlying random quantities, that would rate the quality of the control process. The *closed-loop* information processing strategy is applied, i.e. besides the past observations also the future observation program is taken into account as well.

The cost function $\mathcal{L}_i$ is defined to be quadratic as

$$\mathcal{L}_i = (y_{i+1} - r_{i+1})^2 + qu_i, \tag{12}$$

where r_{i+1} is a known reference signal at time instant $i+1$ and $q > 0$ is a weighting design parameter.

The formal optimal control is described by the backward recursive equation

$$u_k = \underset{u_k}{\operatorname{argmin}} \left\{ E\left\{ \mathcal{L}_k(y_{k+1}, u_k, \boldsymbol{\Theta}_k) + \mathcal{V}_1 \right\} \middle| \boldsymbol{I}^k \right\}, \tag{13}$$

where the scalar function

$$\mathcal{V}_j = \min_{u_i} \left\{ E\left\{ \mathcal{L}_i(y_{i+1}, u_i, \boldsymbol{\Theta}_i) \; + \mathcal{V}_{j+1} \right\} \middle| \boldsymbol{I}^i, u_i \right\} \quad \text{for } i = k+j, \; j = N-1, \ldots, 0 \tag{14}$$

is usually called the Bellman function expressing a minimum average cost-to-go at time step k given the information state $\boldsymbol{I}^k$. The initial condition for the backward recursive equation is usually $\mathcal{V}_N = 0$.

The BOR provides a general tool for finding the optimal control law (13) as a by-product in the process of solving the nested optimisation problem (14). However, the solution is not practically feasible because of previously mentioned reasons and a suboptimal solution has to be sought.

3. CONTROL DESIGN

This section will cover the IMDUC design. First, an approximation of the optimisation problem (10)– (12) is used to obtain a suboptimal feasible solution. An approximation based on the receding horizon technique [Flídr and Šimandl, 2013] is used in this paper because of its simplicity. The length of the receding horizon is shortened as much as possible to reduce computational burden, but on the other hand, to ensure the dual control properties. Then, the problem of evaluation of the necessary expectations is tackled by employing an approximation based on the stochastic integration rule. Finally, the implicit dual control is algorithmically summarised.

3.1 Solution of the optimisation recursion

The first step that makes the BOR more feasible is a reduction of the complexity of the problem. The l-step lookahead policy which approximates the true Bellman function by a heuristically computed one will be employed. The simplest approximation technique is the receding horizon consisting of approximating the cost-to-go at time step $k+l+1$ and beyond by zero. The number of steps $l > 0$ of the limited lookahead policy should be chosen as short as possible. The simplest choice $l = 1$ leads to a minimisation of only the current expected cost at time instant k. However, this would result to the cautious control strategy only, which could generate a very small control signal when the variance of unknown parameters becomes large. In order to ensure both aspects of the dual control, i.e. also the probing feature, it is necessary to solve at least a two-stage optimisation problem [Bar-Shalom and Tse, 1976].

Thus, the aim of the modified optimisation problem with receding horizon reduced to two stages is to find a control law given by (13), where $\mathcal{V}_j = 0$ for $j = N-1, \ldots, 2$, i.e. the possible cost incurred by the control actions u_i, $i > k+1$ is at time k considered as unimportant.

Initially, the expectation of the cost function $\mathcal{L}_i(y_{i+1}, u_i, \boldsymbol{\Theta}_i)$ for $i = 1, 2$ will be examined. Given the quadratic cost function (12), the expectation for system modelled by (2)–(7) can be expressed as

$$\begin{aligned} &E\left\{ (y_{i+1} - r_{i+1})^2 + qu_i \middle| u_i, \boldsymbol{I}^i \right\} = E\left\{ (\hat{f}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) + \hat{g}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) + e_{i+1} \right. \\ &\left. -r_{i+1})^2 \middle| u_i, \boldsymbol{I}^i \right\} = M_i^{ff} + (\mu - r_{i+1})^2 + 2M_i^f(\mu - r_{i+1}) \\ &+ 2(M_i^{fg} + 2M_i^g(\mu - r_{i+1}))u_i + u_i^2(M_i^{gg} + q)^2, \end{aligned} \tag{15}$$

where M_i^{ff}, M_i^{gg}, M_i^{fg}, M_i^f and M_i^g denote following expectations

$$M_i^{ff} = E\left\{ m_i^{ff} \middle| \boldsymbol{I}^i \right\} = E\left\{ \hat{f}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i)\hat{f}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) \middle| \boldsymbol{I}^i \right\}, \tag{16}$$

$$M_i^{gg} = E\left\{ m_i^{gg} \middle| \boldsymbol{I}^i \right\} = E\left\{ \hat{g}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i)\hat{g}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) \middle| \boldsymbol{I}^i \right\}, \tag{17}$$

$$M_i^{fg} = E\left\{ m_i^{fg} \middle| \boldsymbol{I}^i \right\} = E\left\{ \hat{f}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i)\hat{g}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) \middle| \boldsymbol{I}^i \right\}, \tag{18}$$

$$M_i^{f} = E\left\{ m_i^{f} \middle| \boldsymbol{I}^i \right\} = E\left\{ \hat{f}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) \middle| \boldsymbol{I}^i \right\}, \tag{19}$$

$$M_i^{g} = E\left\{ m_i^{g} \middle| \boldsymbol{I}^i \right\} = E\left\{ \hat{g}(\boldsymbol{\Theta}_i, \boldsymbol{x}_i) \middle| \boldsymbol{I}^i \right\}. \tag{20}$$

Even with the parameter conditional pdf $p(\boldsymbol{\Theta}_i | \boldsymbol{I}^i)$ being available, the evaluation of the expectation (16)–(18) is not trivial. Moreover, for $i > k$ the conditional density can not be easily obtained due to the interwovenness of the control and estimation problems. The problem of evaluation of the necessary expectations is the subject of Section 3.2.

In the next step, Equations (15)–(20) can be used for the evaluation of the cost (13). Under the assumption $\mathcal{V}_2 = 0$, the Bellman function $\mathcal{V}_1$ is defined by the relation

$$\mathcal{V}_1 = \min_{u_{k+1}} \left\{ E\left\{ \mathcal{L}_{k+1}(y_{k+2}, u_{k+1}, \boldsymbol{\Theta}_{k+1}) \right\} \middle| u_{k+1}, \boldsymbol{I}^{k+1} \right\} \tag{21}$$

and its value can be determined according to

$$\begin{aligned} \mathcal{V}_1 = \min_{u_{k+1}} \Big\{ & M_{k+1}^{ff} + (\mu - r_{k+2})^2 + 2M_{k+1}^f(\mu - r_{k+2}) \\ & + 2(M_{k+1}^{fg} + M_{k+1}^g(\mu - r_{k+2}))u_{k+1} + u_{k+1}^2(M_{k+1}^{gg} + q)^2 \Big\}. \end{aligned} \tag{22}$$

From Equation (15) it holds that

$$u_{k+1} = -(M_{k+1}^{fg} + M_{k+1}^g(\mu - r_{k+2}))(M_{k+1}^{gg} + q)^{-1}. \tag{23}$$

Then, Equation (22) using (23) can be further revised as

$$\begin{aligned} \mathcal{V}_1 = & M_{k+1}^{ff} + (\mu - r_{k+2})^2 + 2M_{k+1}^f(\mu - r_{k+2}) - (M_{k+1}^{ff})^2 \\ & \cdot (M_{k+1}^{gg} + q)^{-1} - 2M_{k+1}^{fg}M_{k+1}^g(\mu - r_{k+2})(M_{k+1}^{gg} + q)^{-1} \\ & + (M_{k+1}^g)^2(\mu - r_{k+2})^2(M_{k+1}^{gg} + q)^{-1}. \end{aligned} \tag{24}$$

Equation (24) does not explicitly depend on the control u_{k+1} and hence the minimisation operator can be omitted. It is only implicitly dependent on the control u_k through the model output

y_{k+1} and $\boldsymbol{x}_{k+1}$ occurring in the expectations M^{ff}_{k+1}, M^{gg}_{k+1}, M^{fg}_{k+1}, M^{f}_{k+1} and M^{g}_{k+1}. However, its value depends on the control u_k in a nonlinear manner. It means that Equation (13) represents a nonlinear optimisation problem and it is not possible obtain a control law in a closed-form. Therefore, it is inevitable to resort to a numerical solution.

3.2 *Stochastic integration rule*

This section is focused on evaluating the expectations (16)–(20) for the time instants $i = k, k+1$ which are requisite for minimising the control which meets the condition (13). It is necessary to find a suitable approximation that will make the expectation evaluation possible. The usual way is to combine parameter estimation and deterministic control, where the unknown parameters $\boldsymbol{\Theta}_i$ are replaced by their expectations. This heuristically certainty equivalence (HCE) principle ignores potentially significant parameter uncertainties, leading to severe problems such as the bursting phenomenon. Another possibility is to use the Taylor series the first-order approximations for the evaluation [Král and Šimandl, 2011b] but accuracy of the obtained expectations may be insufficient. Hence, an alternative way is to use a numerical approximation of the expectation value calculation. In the paper, a concept based on the stochastic integration rule (SIR) [Genz and J.Monahan, 1998] will be presented. The SIR can be interpreted as a numerical method for enumeration of the expectation function. It is an approximation technique based on a cubature rule evaluated at randomly generated points and thus its result is also random, but with an important property that the result is asymptotically exact.

To find (16)–(20) using the SIR is advisable to introduce the following nonlinear functional

$$\boldsymbol{\varphi}(\boldsymbol{\Theta}_i) = \boldsymbol{I}_5 \times \left[m^{ff}_i, m^{gg}_i, m^{fg}_i, m^{f}_i, m^{g}_i \right]^T, \tag{25}$$

where $\boldsymbol{I}_a$ denotes an $(a \times a)$ identity matrix. This definition has the effect that all of the randomly generated points are common in evaluation of the individual expectations resulting in a significant computational savings.

The expectation μ of a general nonlinear vector function $\boldsymbol{\varphi}(\boldsymbol{\Theta}_i)$, where $\boldsymbol{\Theta}_i$ is a random variable with the normal distribution in the form

$$p(\boldsymbol{\Theta}_i) = \mathcal{N}(\hat{\boldsymbol{\Theta}}_i, \boldsymbol{S}_i\boldsymbol{S}_i^T), \tag{26}$$

where $\boldsymbol{S}_i$ denotes a square root matrix of the covariance $\boldsymbol{P}_i$ such that $\boldsymbol{S}_i\boldsymbol{S}_i^T = \boldsymbol{P}_i$, can be defined by the integral

$$\boldsymbol{\mu}_i = E\left\{\boldsymbol{\varphi}(\boldsymbol{\Theta}_i)\right\} = \int_{\mathbb{R}^{n_\Theta}} \boldsymbol{\varphi}(\boldsymbol{\Theta}_i) p(\boldsymbol{\Theta}_i) d\boldsymbol{\Theta}_i, \tag{27}$$

where $\boldsymbol{\mu}_i = \left[M^{ff}_i, M^{gg}_i, M^{fg}_i, M^{f}_i, M^{g}_i \right]^T$.

Using the substitutions (25)–(27), the SIR proposed for the solution of (16)–(20) is given by the following algorithm:

Algorithm 1: *Stochastic Integration Rule Algorithm*

Step 1: Choose a maximum number of iterations N_{max}.
Step 2: Set the number of iterations $N_I = 0$, initial value of the integral $\tilde{\boldsymbol{\mu}}_i = \boldsymbol{0}$ and compute the vector $\chi_0 = \boldsymbol{\varphi}(\hat{\boldsymbol{\Theta}}_i)$.
Step 3: Repeat (until $N_I < N_{max}$) the following loop:
- **a)** Set $N_I = N_I + 1$.
- **b)** Generate an uniformly random orthogonal matrix $\boldsymbol{U} \in \mathbb{R}^{n_\Theta \times n_\Theta}$ and generate a random number ρ from the χ distribution, i.e. $\rho \sim \chi(n_\Theta + 2)$.
- **c)** Compute a set of vectors χ_j for $j = 1, \ldots, n_\Theta$ according to

$$\chi_j = -\rho \boldsymbol{S}_i \boldsymbol{U} \boldsymbol{e}_j + \hat{\boldsymbol{\Theta}}_i, \tag{28}$$

$$\chi_{n_\Theta + j} = \rho \boldsymbol{S}_i \boldsymbol{U} \boldsymbol{e}_j + \hat{\boldsymbol{\Theta}}_i, \tag{29}$$

where $\boldsymbol{e}_j$ is the j-th column of the identity matrix $\boldsymbol{I}_{n_\Theta}$, and the corresponding weights ω_j as

$$\omega_0 = \frac{n_\Theta}{\rho^2}, \tag{30}$$

$$\omega_j = \omega_{n_\Theta + j} = \frac{n_\Theta}{2\rho^2}. \tag{31}$$

- **d)** Compute the value Q of the integral at current iteration

$$Q = \sum_{j=0}^{2n_\Theta} \boldsymbol{\varphi}(\chi_j)\omega_j \tag{32}$$

and use it to update the approximate value $\tilde{\boldsymbol{\mu}}$

$$\tilde{\boldsymbol{\mu}}_i = \tilde{\boldsymbol{\mu}}_i + (Q - \tilde{\boldsymbol{\mu}}_i)/N_I. \tag{33}$$

Step 4: The approximate value of the integral $\boldsymbol{\mu}_i$ is given by $\tilde{\boldsymbol{\mu}}_i$.

3.3 *Implicit dual control algorithm*

The solution to the optimisation recursion suggested in Section 3.1 accompanied by the SIR method for evaluation on the expectations (16)–(20) presented in Section 3.2 constitutes a basis of the IMDUC algorithm proposed in this subsection.

The aim of the algorithm is to describe the technique for calculation of the dual control law defined by (13). Unfortunately, a closed-form formula cannot be derived, that would define the control minimising the inner expectation $\left\{E\left\{\mathcal{L}_k(y_{k+1}, u_k, \boldsymbol{\Theta}_k) + \mathcal{V}_1\right\} \middle| u_k, \boldsymbol{I}_k\right\}$. Thus, it is unavoidable to seek the minimising control u_k using a numerical optimisation method. The following algorithm describes the main steps for generating action signal u_k at time instant k:

Algorithm 2: *Implicit Dual Control Algorithm*

Step 1: Obtain the measurement of the system output y_k.
Step 2: Determine the mean $\hat{\boldsymbol{\Theta}}_k$ and the covariance matrix $\boldsymbol{P}_k$ of the unknown model parameters using the EKF as the estimator.
Step 3: Set the iteration counter $\ell = 0$ and choose an initial candidate for the suboptimal control $u_k^{(\ell)}$.
Step 4: Repeat the following loop until a satisfactory minimising control u_k is found:
- **a)** Determine the prediction of the system output

$$y_{k+1} = \hat{f}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k) + \hat{g}(\boldsymbol{\Theta}_k, \boldsymbol{x}_k) u_k^{(\ell)}. \tag{34}$$

- **b)** Calculate the approximation of Bellman function $\mathcal{V}_1^{(\ell)}$ using Equation (24).
- **c)** Evaluate the cost-to-go of the modified optimisation problem stated in Section 3.1

$$J_k^{(\ell)} = \left\{E\left\{\mathcal{L}_k^{(\ell)}(y_{k+1}, u_k^{(\ell)}, \boldsymbol{\Theta}_k) + \mathcal{V}_1^{(\ell)}\right\} \middle| u_k^{(\ell)}, \boldsymbol{I}_k\right\}. \tag{35}$$

- **d)** Test if the relative change of the cost-to-go $J_k^{(\ell)}$ is satisfactory low. If it is not, increment the counter ℓ, choose a new control candidate $u_k^{(\ell)}$ and proceed with Step 4a).

The NN parameter estimation in Steps 1 can be determined by employing any suitable nonlinear filtering technique. The promising method seems to be a novel filter based on SIR as well [Duník et al., 2013]. Although computational demands are by several orders higher than the EKF, the benefit of this choice besides a good quality of the estimates would be in the possibility to reuse the points generated by the SIR. Nevertheless, an analysis of another estimation method utilization remains an open problem.

The initial candidate for the suboptimal control $u_k^{(0)}$ in Step 3 can be determined using for example the cautious, HCE or some explicit dual controller. The SIR approximation described in Section 3.2 is employed in Steps 4b) and 4c) where it is used to evaluate the expectations (16)–(20) repeatedly. Finally, the minimisation of the cost-to-go in Step 4 is accomplished using a suitable numerical optimisation method.

4. NUMERICAL EXAMPLE

Properties of the designed IMDUC are tested and validated in the following numerical example

$$y_k = \frac{-1.5y_{k-1}y_{k-2}}{1+y_{k-1}^2+y_{k-2}^2} + 0.35\sin(y_{k-1}+y_{k-2}) + (2+\cos(y_{k-1}y_{k-2}))u_{k-1} + e_k, \tag{36}$$

where $n_y = 2$, $n_u = 0$, the sequence $\{e_k\}$ is a white noise with zero mean and variance $\sigma^2 = 0.001$. System (36) is approximated by the model described in Equations (2)–(7) structured with $nf = 10$ and $ng = 5$ neurons, resulting in a total of $N = 62$ parameters for nonlinear estimation. The initial condition $\boldsymbol{x}_0$ is considered to be $\mathcal{N}\{0,\ 0.001\boldsymbol{I}_{n_y+n_u}\}$ and the initial vector $\hat{\boldsymbol{\Theta}}_0$ of the nonlinear model is generated from the normal distribution $\mathcal{N}\{0,\ 0.1\boldsymbol{I}_{n_\Theta}\}$, the covariance matrix should reflect confidence in the initial guess and is chosen as $\boldsymbol{P}_0 = 0.1\boldsymbol{I}_{n_\Theta}$. The reference signal r_k was chosen as a square wave with unit amplitude and period 25 time instants. Further, the EKF is used as a parameter estimator and the cautious control was chosen to be convenient for initialization of $u_k^{(0)}$.

It must be emphasized, that there is no off-line training of the NN based model. Thus, the uncertainty in the model is involved, both due to the external disturbances e_k and due to imprecise knowledge of the controlled system (i.e. the unknown functions $f(\boldsymbol{x}_k)$ a $g(\boldsymbol{x}_k)$).

The proposed implicit adaptive dual control is compared with a couple of explicit solutions; the innovation dual control (IDC) and bicriterial dual controller (BDC) which were presented in Fabri and Kadirkamanathan [1998] and Šimandl et al. [2005], respectively. Both explicit controllers include several tuning parameters, which were selected by trial-and-error to obtain a minimal value from a chosen criteria point of view.

The quality of control will be evaluated by M Monte Carlo simulations. The function $\mathcal{L}$ is chosen as an accumulated cost

$$\mathcal{L} = \sum_{k=1}^{N-1}(y_{k+1} - r_{k+1})^2 + 0.01u_k^2,$$

The value of the cost $\mathcal{L}$ for the particular j^{th} Monte Carlo simulation is denoted by $\mathcal{L}_j$. The criterion J is estimated by its mean $\hat{J} = \frac{1}{M}\sum_{j=1}^{M}\mathcal{L}_j$ and median of $\mathcal{L}$ represented by $\tilde{J}$. Variability of Monte Carlo simulations is expressed by

$$\text{var}\{\mathcal{L}\} = \frac{1}{M-1}\sum_{j=1}^{M}(\mathcal{L}_j - \hat{J})^2$$

and the quality of the criterion estimate $\hat{J}$ is represented by var$\{\hat{J}\}$ which can be computed using the bootstrap technique [Efron and Tibshirani, 1994].

The criterion value estimates ($\hat{J}$, $\tilde{J}$), the accuracy of the estimate (var$\{\hat{J}\}$), the variability of Monte Carlo simulations (var$\{\mathcal{L}\}$), and the average time per step that were computed using $M = 200$ Monte Carlo simulations with $N = 100$ steps per simulation, are summarised in Table I. Although the behaviour in each of the three schemes is satisfactory, the best statistical measures were achieved for the IMDUC. The lowest values of $\hat{J}$, $\tilde{J}$, var$\{\hat{J}\}$ and var$\{\mathcal{L}\}$ indicate superior control performance which is illustrated by Figure 1, where a typical realization of the system controlled by the IMDUC is plotted. Thus, the proposed solution qualitatively indicates the best compromise between the conflicting objectives of dual control, i.e. caution and probing. Additionally, there is no need to tune any additional design parameters. A notable drawback is a higher computational burden compared to the explicit ones due to the use of the numerical approximate methods.

TABLE I. A quality control performance of the BDC, IDC and IMDUC based on the Monte Carlo simulations.

Controller	$\hat{J}$	$\tilde{J}$	var$\{\mathcal{L}\}$	var$\{\hat{J}\}$	time[s]
IDC	28.88	26.99	168.17	1.65	0.0056
BDC	23.86	21.94	106.04	0.55	0.0056
IMDUC	17.99	15.98	72.62	0.26	3.5

The results of intensive numerical simulations are depicted in Figure 2, where the minimum and maximum as the extremes values of the output at individual time instants are illustrated. The time developments, which are based on the Monte-Carlo simulations, represent imaginary envelopes of the worst-case behaviour for all three controller configurations. It can be seen that the performance of the IMDUC qualitatively shows the best transient performance and the most smooth response, avoiding excessive overshoots. It can be also mentioned that the difference among the output responses vanishes after time instant $k = 80$ as the system uncertainty is reduced.

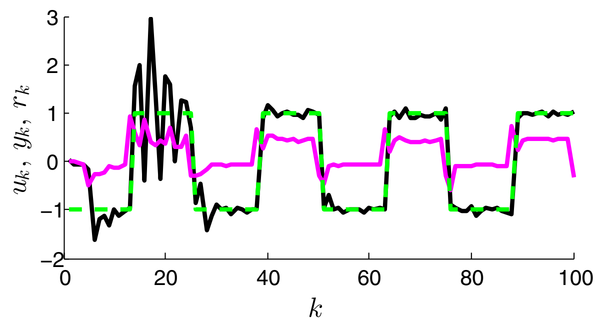

Fig. 1. The typical output response (black), input (magenta) and reference signal (green) of a single trial for the system controlled by the IMDUC.

5. CONCLUSION

Implicit type of dual control was proposed for a class of nonlinear stochastic systems subject to functional uncertainty. The underlying idea of the control design is based on several approximations to avoid the problem with a closed-form solvability of the Bellman optimisation recursion. An infinite horizon constrained optimisation problem was tackled by reformulation

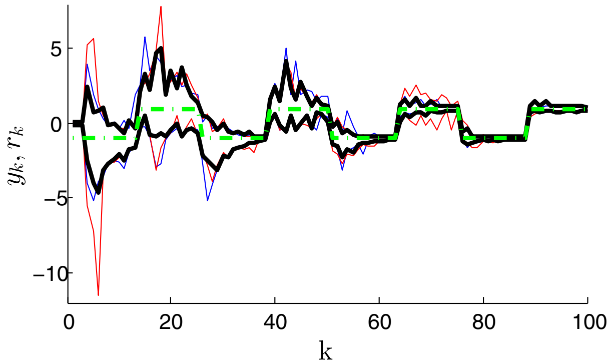

Fig. 2. The chosen reference signal (green) and the extreme output values at the individual time instants based on the Monte-Carlo simulations for the IDC (red), BDC (blue) and IMDUC (black).

to the receding horizon type of the optimisation problem and by reduction to a two-stage optimisation. The inherent obstacle of determining the expectation was eliminated by employing a powerful technique for approximate evaluation of expectations based on the stochastic integration rule. Finally, a numerical algorithm for calculation of the control action was presented. This successfully generalises the work proposed by Flídr and Šimandl [2013] to a generic class of nonlinear systems.

The performance of the designed controller was evaluated through extensive numerical simulations followed by a Monte Carlo analysis. It was shown that the proposed solution is not only capable superior to the current solutions of functional adaptive control, but it also eliminates their drawback of tuning additional design parameters. The proposed concept accounts for a higher computational burden, but at the same time shows an excellent control performance, thus provide a promising alternative to the existing, explicit type based, dual controllers.

Further research could be focused twofold: (i) to reduce the computational burden, e.g. by a parallel computing, or (ii) to analyse a utilization of other estimation methods as indicated in Section 3.3.

REFERENCES

Y. Bar-Shalom and E. T. S. Tse. Concepts and methods in stochastic control. *Control and Dynamic Systems*, 12:99–172, 1976.

D. Bayard and A. Schumitzky. Implicit dual control based on particle filtering and forward dynamic programming. *Int. J. Adapt. Control Signal Process*, 24(3):155–177, 2010.

M. K. Bugeja, S. G. Fabri, and L. Camilleri. Dual adaptive control of mobile robots using neural networks. *IEEE Transactions on Systems, Man and Cybernetics - Part B*, 39 (1):129–141, 2009.

Fu-Chuang Chen and H. K. Khalil. Adaptive control of a class of nonlinear discrete-time systems using neural networks. *IEEE Trans. on Automatic Control*, 40(5):791–801, 1995.

J. Duník, M., Šimandl, and O. Straka. Stochastic integration filter. *IEEE Transactions on Automatic Control*, 58(6):1561–1566, 2013.

B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman and Hall, 1994.

S. G. Fabri and M. K. Bugeja. Unscented transform-based dual adaptive control of nonlinear MIMO systems. In *Proceedings of* 13^{th} *European Control Conference*, pages 392–397, 2013. Zürich, Switzerland.

S. G. Fabri and V. Kadirkamanathan. Dual adaptive control of nonlinear stochastic systems using neural networks. *Automatica*, 34(2):245–253, 1998.

S. G. Fabri and V. Kadirkamanathan. *Functional Adaptive Control*. Springer-Verlag, 2001.

A. A. Fel'dbaum. *Optimal Control Systems*. Academic Press, New York, 1965.

N. M. Filatov and H. Unbehauen. *Adaptive Dual Control*. Springer-Verlag, 2004.

M. Flídr and M. Šimandl. Implicit dual controller based on stochastic integration rule. In *Proceedings of the* 13^{th} *European Control Conference*. Zürich, Switzerland, 2013.

A. Genz and J.Monahan. Stochastic integration rules for infinite regions. *SIAM J. Sci. Comput.*, 19(2):426–439, 1998.

S. Haykin. *Kalman Filtering and Neural Networks*. Wiley-Interscience, 2001.

S. Haykin. *Neural Networks: A comprehensive foundation*. Prentice-Hall, Upper Saddle River, NJ, 2nd edition, 1999.

R. Herzallah and D. Lowe. A novel approach to modelling and exploiting uncertainty in stochastic control systems. In *Artificial Neural Networks - ICANN 2002*, pages 801–806. Springer Berlin Heidelberg, 2002.

R. Herzallah and D. Lowe. A bayesian perspective on stochastic neurocontrol. *IEEE Transactions on Neural Networks*, 19(5): 914–924, 2008.

L. Král and M. Šimandl. Predictive dual control for nonlinear stochastic systems modelled by neural networks. In *Proceedings of the 19th Mediterranean Conference on Control and Automation*. Corfu, Greece, 2011a.

L. Král and M. Šimandl. Dual adaptive control for non-minimum phase systems with functional uncertainties. In *Proceedings of* 9^{th} *IFAC Symposium on Nonlinear Control Systems*, pages 515–520. Toulouse, France, 2013.

L. Král and M. Šimandl. Functional adaptive controller for multivariable stochastic systems with dynamic structure of neural network. *Int. J. Adapt. Control Signal Process.*, 25: 949–964, 2011b.

J. M. Lee and J. H. Lee. An approximate dynamic programming based approach to dual adaptive control. *Journal of Process Control*, 19(5):859 – 864, 2009.

R. Milito, C. S. Padilla, R. A. Padilla, and D. Cadorin. An inovations approach to dual control. *IEEE Trans. on Automatic Control*, 27(1):133–137, 1982.

J. Sarangapani. *Neural Network Control of Nonlinear Discrete-Time Systems*. Taylor & Francis, 2006.

E. Tse, Y. Bar-Shalom, and L. Meier. Wide-sense adaptive dual control for nonlinear stochastic systems. *IEEE Transactions on Automatic Control*, 18(2):98–108, 1973.

M. Šimandl, L. Král, and P. Hering. Neural network based bicriterial dual control of nonlinear systems. In *Proceedings of the* 16^{th} *IFAC World Congress*. Prague: IFAC, 2005.

J. Wieselander and B. Wittenmark. An approach to adaptive control using real-time identification. *Automatica*, 7(2):211–217, 1971.