

Implicit Dual Controller based on Stochastic Integration Rule

Miroslav Flidr and Miroslav Šimandl

Abstract—A new implicit dual control method is proposed providing a suboptimal solution of the optimal control problem for discrete linear stochastic state space model with unknown and unobservable parameters. The solution is based on Bellman optimization recursion where two stages of optimization recursion will be pursued. The resulting controller ensures both dual properties of the suboptimal control, i.e. caution and probing. In order to be able to determine the control, the stochastic integration rule is employed for approximate evaluation of expectations. The dual control is then obtained using suitable iterative numerical algorithm. The proposed implicit dual controller is compared to the explicit dual controllers which are easier to derive but require proper tuning of design parameters.

I. INTRODUCTION

The solution to the stochastic optimal control problem is known to be formally attained by employing the Bellman optimization recursion (BOR). Unfortunately, the derivation of the control law is often prohibitive. The designer is faced with another equally tedious problem of nonlinear estimation, i.e. the necessity to solve the Bayesian recursive relations (BRR). Moreover, the attempt to make use of a numerical approach is hindered by the so called curse of dimensionality. As was pointed out by Feldbaum [1], the result of the aforementioned interwovenness of control and estimation is the exhibition of two conflicting features of the control. The control, referred to as dual control, beside the fulfillment of the control objective also actively influences the reduction of uncertainty in knowledge of the unobserved states and parameters. Because of the difficulties with finding the optimal solution various simplifications of the original problem are used leading to a suboptimal control.

The closed form solution is attainable only in a few special cases. The most notable special case is the problem of optimal control of a linear Gaussian system with quadratic criterion (LQG). Because of the clarity and simplicity of the LQG control law it is often used also for related problem where the system depends on unknown and unobservable parameters. The resulting control law is often denoted as the heuristic certainty equivalence controller (HCE) because the inherent uncertainties in the knowledge of the unobserved parameters are not taken into account and the certainty equivalence property is enforced on the problem. Due to the total omission of the uncertainties, the HCE controller may cause inappropriate behavior because the controller blindly

uses the estimated parameters as the true ones even in case where there is high uncertainty in these estimates. Another option is to use a cautious type of controller [2], [3]. The cautious control law is relatively easy to design and does respect the inherent uncertainties. Its deficiency is that it can be prone to the so called turn-off phenomenon [4] and it cannot actively help with the uncertainty reduction.

The deficiencies of both HCE and cautious controllers drive the ongoing effort to develop methods that would lead to a suboptimal solution preserving both features of the dual control. These methods can be classified as *explicit* and *implicit* ones [5]. The *explicit* type dual control methods can be seen as more simple and straightforward. They employ a simplification of the original optimization problem in order to make the closed form solution possible and a suitable extension reassuring the appearance of measure of the uncertainty in the problem. The simplification itself mostly guarantees the cautious property only and the extension induces the probing property of the resulting control [6], [7], [8]. The *implicit* type of dual control methods employ various approximation techniques approximating either the Bellman function or the probability density functions by retaining the dual features of the control [9], [10]. The wider application of these approaches is hindered by the fact that they are too complicated for practical implementation and have very high computational demands.

One possible way how to tackle the design of implicit dual controller is to suitably approximate the evaluation of the expectations occurring in the BOR. In essence it is necessary to choose a computationally effective way to approximate the integral evaluation. Besides the Monte Carlo (MC) integration the stochastic integration rule [11], [12] can be used. This rule provide an asymptotically exact integral evaluation with faster convergence than the MC integration and possibly with significantly lower computational costs. In order to make the design computationally tractable the problem can be reformulated to the receding horizon type.

The goal of the paper is to propose an algorithm of novel implicit dual controller which naturally comprises both aspects of the dual control. The presentation will encompass discussion of the BOR solvability, simplification of the whole controller design to the receding horizon type with two-step ahead horizon and choice of approximate expectation calculation.

The article is organized as follows. In the next section the problem of optimal control of the linear discrete stochastic system with unknown unobservable parameters will be stated. In Section III the general solution in the form of the BOR will be presented and the obstacles for finding

This work was supported by the Czech Science Foundation, project GACR P103/11/1353.

M. Flidr and M. Šimandl are with the Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia, Univerzitní 8, 30614 Plzeň, Czech Republic, e-mail: flidr@kky.zcu.cz, simandl@kky.zcu.cz

the closed form control law will be pointed out. Then, in Section IV the employed approximation technique and the algorithm for control calculation will be presented. Finally, the numerical example will be presented in Section V.

II. OPTIMAL CONTROL PROBLEM FORMULATION

The goal of the optimization problem stated in this section is to find the optimal control minimizing the classical quadratic criterion subject to discrete time stochastic state space system with unknown parameters.

Consider the discrete time stochastic system given by the following relations

$$\mathbf{s}_{k+1} = \mathbf{A}(\boldsymbol{\theta}_k)\mathbf{s}_k + \mathbf{B}(\boldsymbol{\theta}_k)\mathbf{u}_k + \mathbf{w}_k, \quad (1)$$

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\Phi}_k\boldsymbol{\theta}_k + \boldsymbol{\epsilon}_k, \quad k = 0, \dots, N-1 \quad (2)$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{s}_k) + \mathbf{v}_k \quad (3)$$

describing the time evolution of the state (1) and the parameters (2) and equation specifying the measurement (3). The vector $\mathbf{s}_k \in \mathbb{R}^n$ represents the state of the controlled system, $\boldsymbol{\theta}_k \in \mathbb{R}^p$ the vector of the unknown parameters, $\mathbf{y}_k \in \mathbb{R}^m$ the measurement vector of the controlled system and $\mathbf{u}_k \in \mathbb{R}^r$ is the control input vector. The quantities $\mathbf{w}_k \in \mathbb{R}^n$, $\boldsymbol{\epsilon}_k \in \mathbb{R}^p$, $\mathbf{v}_k \in \mathbb{R}^m$ are state, parameter and measurement noises, respectively. The random processes $\{\mathbf{w}_k\}$, $\{\boldsymbol{\epsilon}_k\}$ and $\{\mathbf{v}_k\}$ are white, mutually independent sequences and are independent of the random quantities \mathbf{s}_0 and $\boldsymbol{\theta}_0$ as well. The probability density functions (pdfs) of the random quantities \mathbf{s}_0 , $\boldsymbol{\theta}_0$, \mathbf{w}_k , $\boldsymbol{\epsilon}_k$, \mathbf{v}_k are supposed to be known. The elements of the matrices $\mathbf{A}(\boldsymbol{\theta}_k)$ and $\mathbf{B}(\boldsymbol{\theta}_k)$ are known linear functions of the unknown random parameters $\boldsymbol{\theta}_k$. The dimensions of the matrices $\mathbf{A}(\boldsymbol{\theta}_k)$ and $\mathbf{B}(\boldsymbol{\theta}_k)$ correspond with the dimensions of the state vector \mathbf{s}_k and the control input \mathbf{u}_k , respectively. The functions $\mathbf{h}_k(\mathbf{s}_k)$ and matrices $\boldsymbol{\Phi}_k$ are completely known and have appropriate dimensions.

The aim is to find the control law

$$\mathbf{u}_k = \mathbf{u}_k(\mathcal{J}_k), \quad k = 0, 1, \dots, N-1 \quad (4)$$

that minimizes criterion

$$J = E \left\{ \sum_{k=0}^{N-1} \mathcal{L}_k(\mathbf{s}_k, \boldsymbol{\theta}_k, \mathbf{u}_k) \right\}, \quad (5)$$

subject to the discrete time stochastic MIMO system given by relations (1)–(3). The expectation $E\{\cdot\}$ is taken over all the underlying random quantities. The *closed-loop* information processing strategy is considered, i.e. besides the past observations also the future observation program is taken into account.

The symbol \mathcal{J}_k represents complete information available up to the time instant k , i.e.

$$\mathcal{J}_k = (\mathbf{u}_0^{k-1}, \mathbf{y}_0^k) \quad (6)$$

where $\mathbf{u}_0^{N-1} \triangleq (\mathbf{u}_0, \dots, \mathbf{u}_{N-1})$ and $\mathbf{y}_0^{N-1} \triangleq (\mathbf{y}_0, \dots, \mathbf{y}_{N-1})$. The cost function is considered to be quadratic and is defined as

$$\mathcal{L}_k(\cdot) = (\mathbf{s}_{k+1} - \bar{\mathbf{s}}_{k+1})^T \mathbf{Q}_{k+1} (\mathbf{s}_{k+1} - \bar{\mathbf{s}}_{k+1}) + \mathbf{u}_k^T \mathbf{R}_k \mathbf{u}_k, \quad (7)$$

The quantity $\bar{\mathbf{s}}_{k+1} \in \mathbb{R}^n$ denotes the setpoint vector at time instant $k+1$. The state \mathbf{s}_{k+1} depends on \mathbf{s}_k and $\boldsymbol{\theta}_k$ according to (1). The matrices $\mathbf{Q}_{k+1} \in \mathbb{R}^{n \times n}$ and $\mathbf{R}_k \in \mathbb{R}^{r \times r}$ are appropriately chosen positive semidefinite and positive definite matrices, respectively.

Even though this problem may, on the first sight, imply easy solvability, evoked by the resemblance to the well known LQG problem, the opposite is actually true. The closed-form solution to this optimization problem is not attainable. This is brought about by the uncertainty in the knowledge of the parameters. Even if the measurement equation (3) would be linear, the system is in fact nonlinear from the estimation point of view making it harder to achieve the solution. Both the estimation and optimal control task are tightly interwoven in this problem. This interweaving is actually the natural cause that leads to control law with dual properties.

III. GENERAL OPTIMIZATION PROBLEM SOLUTION

This section will cover the general solution to the optimization problem stated in the previous section. The problem will be tackled by means of the BOR. Because of adaptive nature of the problem given by the unobservability of the system parameters, it is essential to reflect the possibility that the estimated parameters may vary at each time step. Thus, the problem will be firstly reformulated as the receding horizon type of the optimization problem, i.e. only the suboptimal solution to the original optimization problem will be sought.

The new problem is stated such as at each time step the optimal control law is sought minimizing the cost-to-go criterion

$$\mathbf{u}_k = \underset{u_k}{\operatorname{argmin}} J_k(\mathcal{J}_k), \quad k = 0, 1, \dots, N-1 \quad (8)$$

where the cost-to-go criterion is defined as follows

$$J_k(\mathcal{J}_k) = E \left\{ \sum_{i=k}^{k+m} \mathcal{L}_i(\mathbf{s}_i, \boldsymbol{\theta}_i, \mathbf{u}_i) \mid \mathcal{J}_k \right\} \quad (9)$$

with m denoting the length of the so-called receding horizon.

A. Bellman optimization recursion

The BOR provides a general tool for finding the optimal control law. It transforms the complex problem of multistage optimization to multiple nested static optimization problems. In the process of solving these static optimization problems, i.e. by looking for the Bellman function in a particular step, the optimal control law valid at this time step is found as byproduct.

The recursive relation for determining the Bellman functions for all the time instants $i = k, \dots, k+m$ is for the receding horizon type of problem generally defined by the relation

$$\mathcal{V}_j(\mathcal{J}_i) = \min_{u_i} \left\{ E \left\{ \mathcal{L}_i(\mathbf{s}_i, \boldsymbol{\theta}_i, \mathbf{u}_i) + \mathcal{V}_{j+1}(\mathbf{u}_i, \mathcal{J}_i) \right\} \right\}, \quad (10)$$

$i = k+j; j = m, \dots, 0$

where the cost function at time i denoted as $\mathcal{L}_i(s_i, \boldsymbol{\theta}_i, \mathbf{u}_i)$ is defined by (7) and the boundary conditions is zero, i.e.

$$\mathcal{V}_{m+1} = \boldsymbol{\mathcal{O}}. \quad (11)$$

Since the receding horizon type of the optimization problem is pursued, only the control \mathbf{u}_k will be applied. This control minimizes the last step of the optimization recursion (i.e. for $j = 0$) and thus determines the value of the Bellman function $\mathcal{V}_0(\mathcal{J}_k)$. The Bellman function is defined as minimal value of the cost-to-go criterion (9) and thus the minimizing control \mathbf{u}_k inevitably satisfies (8) and the following relations holds

$$\mathbf{u}_k = \underset{u_k}{\operatorname{argmin}} \left\{ E \left\{ \mathcal{L}_k(s_k, \boldsymbol{\theta}_k, \mathbf{u}_k) + \mathcal{V}_1 \left| \mathbf{u}_k, \mathcal{J}_k \right. \right\} \right\}. \quad (12)$$

B. Solvability of Bellman optimization recursion

The BOR does not present the final solution; in fact it only constitutes a recipe on how to find the optimal control law. This section will present the obstacles that prevent the closed form solvability of the problem at hand.

Initially, the influence of the Bellman function \mathcal{V}_j on the minimization in (10) will be put aside and only the expectation of the cost function $\mathcal{L}_i(s_i, \boldsymbol{\theta}_i, \mathbf{u}_i)$ for $i = k, \dots, k+m$ will be examined. Given the quadratic cost function defined in (7) the expectation can be expressed as

$$\begin{aligned} & E \left\{ (s_{i+1} - \bar{s}_{i+1})^T \mathbf{Q}_{i+1} (s_{i+1} - \bar{s}_{i+1}) + \mathbf{u}_i^T \mathbf{R}_i \mathbf{u}_i \left| \mathbf{u}_i, \mathcal{J}_i \right. \right\} = \\ & = M_i^{sAs} + (\hat{\mathbf{w}}_i - \bar{s}_{i+1})^T \mathbf{Q}_{i+1} (\hat{\mathbf{w}}_i - \bar{s}_{i+1}) + \operatorname{tr} \left\{ \mathbf{Q}_{i+1} \mathbf{P}_i^w \right\} \\ & + M_i^{sAB} \mathbf{Q}_{i+1} (\hat{\mathbf{w}}_i - \bar{s}_{i+1}) + (\hat{\mathbf{w}}_i - \bar{s}_{i+1})^T \mathbf{Q}_{i+1} M_i^{As} \\ & + \mathbf{u}_i^T \left(M_i^{BAAs} + \mathbf{B}^T(\hat{\boldsymbol{\theta}}_i) \mathbf{Q}_{i+1} \hat{\mathbf{w}}_i - \mathbf{B}^T(\hat{\boldsymbol{\theta}}_i) \mathbf{Q}_{i+1} \bar{s}_{i+1} \right) \\ & + \left(M_i^{ABAs} + \hat{\mathbf{w}}_i^T \mathbf{Q}_{i+1} \mathbf{B}(\hat{\boldsymbol{\theta}}_i) - \bar{s}_{i+1}^T \mathbf{Q}_{i+1} \mathbf{B}(\hat{\boldsymbol{\theta}}_i) \right) \mathbf{u}_i \\ & + \mathbf{u}_i^T \left(\mathbf{R}_i + M_i^{BB} \right) \mathbf{u}_i, \end{aligned} \quad (13)$$

where $\hat{\mathbf{w}}_i$ and \mathbf{P}_i^w represent state noise mean and covariance matrix, respectively, and M_i^{sAs} , M_i^{sAB} , M_i^{BAAs} , M_i^{As} , M_i^{As} and M_i^{BB} denote following expectations

$$M_i^{sAs} = E \left\{ \mathbf{s}_i^T \mathbf{A}^T(\boldsymbol{\theta}_i) \mathbf{Q}_{i+1} \mathbf{A}(\boldsymbol{\theta}_i) \mathbf{s}_i \left| \mathbf{u}_i, \mathcal{J}_i \right. \right\}, \quad (14)$$

$$M_i^{sAB} = E \left\{ \mathbf{s}_i^T \mathbf{A}^T(\boldsymbol{\theta}_i) \mathbf{Q}_{i+1} \mathbf{B}(\boldsymbol{\theta}_i) \left| \mathbf{u}_i, \mathcal{J}_i \right. \right\}, \quad (15)$$

$$M_i^{BAAs} = \left(M_i^{sAB} \right)^T, \quad (16)$$

$$M_i^{As} = E \left\{ \mathbf{s}_i^T \mathbf{A}^T(\boldsymbol{\theta}_i) \left| \mathbf{u}_i, \mathcal{J}_i \right. \right\}, \quad (17)$$

$$M_i^{As} = \left(M_i^{sA} \right)^T, \quad (18)$$

$$M_i^{BB} = E \left\{ \mathbf{B}^T(\boldsymbol{\theta}_i) \mathbf{Q}_{i+1} \mathbf{B}(\boldsymbol{\theta}_i) \left| \mathbf{u}_i, \mathcal{J}_i \right. \right\}. \quad (19)$$

Even provided that the conditional pdf of the state and parameters $p(s_i, \boldsymbol{\theta}_i | \mathcal{J}_i)$ is available, the evaluation of the expectation (14)-(16) is not trivial. Moreover, for $i > k$ the conditional density is not easily obtainable due to the interwovenness of the control and estimation problems.

Based only on the form of the expectation (13), it would seem that the minimization could be easily accomplished.

However, even though the form of the expectation (13) of the cost function is quadratic with respect to control \mathbf{u}_i , the same cannot be said about the expectation of Bellman function $E \left\{ \mathcal{V}_{j+1} | \mathcal{J}_i \right\}$ with $j = i - k + 1$. Its value depends on the control \mathbf{u}_i in a nonlinear manner.

IV. RECEDING HORIZON IMPLICIT DUAL CONTROLLER

This section will be devoted to description of the novel adaptive dual controller. The basic idea behind the new design is twofold. Firstly, the length of the receding horizon is shortened to make the solving the optimization recursion feasible but still ensuring the dual control properties. Second, the problem of evaluation of the necessary expectations is tackled by employing an approximation based on the stochastic integration rule.

A. Solution of the optimization recursion

The first step that will make the BOR more feasible to solve will be to reduce the length of the supposed receding horizon. This reduces the complexity of the problem. The simplest choice would seem to be minimization of only the current expected cost at time instant k . However, this would lead to the cautious control strategy only. 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 optimization problem.

The aim of the modified optimization problem with receding horizon reduced to two stages is to find a control law given as

$$\mathbf{u}_k = \underset{u_k}{\operatorname{argmin}} \left\{ E \left\{ \mathcal{L}_k(s_k, \boldsymbol{\theta}_k, \mathbf{u}_k) + \mathcal{V}_1 \left| \mathbf{u}_k, \mathcal{J}_k \right. \right\} \right\}, \quad (20)$$

where $\mathcal{V}_j = \boldsymbol{\mathcal{O}}$ for $j = m, \dots, 2$, i.e. the possible cost incurred by the control actions \mathbf{u}_i , $i > k + 1$ is at time k considered as unimportant.

Under the assumption that $\mathcal{V}_2 = 0$, the Bellman function \mathcal{V}_1 is defined by the relation

$$\mathcal{V}_1 = \min_{u_{k+1}} \left\{ E \left\{ \mathcal{L}_{k+1}(s_{k+1}, \boldsymbol{\theta}_{k+1}, \mathbf{u}_{k+1}) \left| \mathbf{u}_{k+1}, \mathcal{J}_{k+1} \right. \right\} \right\} \quad (21)$$

and its value can be determined according to the equation

$$\begin{aligned} \mathcal{V}_1 = & M_{k+1}^{sAs} - M_{k+1}^{sAB} \left(\mathbf{R}_{k+1} + M_{k+1}^{BB} \right)^{-1} M_{k+1}^{BAAs} \\ & + \left(M_{k+1}^{sA} - M_{k+1}^{sAB} \left(\mathbf{R}_{k+1} + M_{k+1}^{BB} \right)^{-1} \mathbf{B}^T(\hat{\boldsymbol{\theta}}_{k+1}) \right) \\ & \times \mathbf{Q}_{k+2} (\hat{\mathbf{w}}_{k+1} - \bar{s}_{k+2}) \\ & + (\hat{\mathbf{w}}_{k+1} - \bar{s}_{k+2})^T \mathbf{Q}_{k+2} \\ & \times \left(M_{k+1}^{As} - \mathbf{B}(\hat{\boldsymbol{\theta}}_{k+1}) \left(\mathbf{R}_{k+1} + M_{k+1}^{BB} \right)^{-1} M_{k+1}^{BAAs} \right) \\ & + (\hat{\mathbf{w}}_{k+1} - \bar{s}_{k+2})^T \mathbf{Q}_{k+2} (\hat{\mathbf{w}}_{k+1} - \bar{s}_{k+2}) \\ & - (\hat{\mathbf{w}}_{k+1} - \bar{s}_{k+2})^T \mathbf{Q}_{k+2} \mathbf{B}(\hat{\boldsymbol{\theta}}_{k+1}) \left(\mathbf{R}_{k+1} + M_{k+1}^{BB} \right)^{-1} \\ & \times \mathbf{B}^T(\hat{\boldsymbol{\theta}}_{k+1}) \mathbf{Q}_{k+2} (\hat{\mathbf{w}}_{k+1} - \bar{s}_{k+2}) + \operatorname{tr} \left\{ \mathbf{Q}_{k+2} \mathbf{P}_{k+1}^w \right\}. \end{aligned} \quad (22)$$

The equation (22) does not explicitly depend on control \mathbf{u}_{k+1} which would not be used anyway. It is only implicitly

dependent on the control \mathbf{u}_k through the state \mathbf{s}_{k+1} occurring in expectations M_{k+1}^{sAs} , M_{k+1}^{sAB} , $M_{k+1}^{BA s}$, M_{k+1}^{sA} and M_{k+1}^{As} .

B. Stochastic integration rule

In order to be able to search for the minimizing control which meets the condition (20), it is necessary to evaluate the expectations (14)-(19) for the time instants $i = k, k + 1$. The evaluation of the expectations (17)-(19) is quite straightforward; however, the remaining expectations (14)-(16) are harder to obtain. It is necessary to resort to a suitable approximation that will make the expectation evaluation possible. One possible way is used for the design of the PCE controller or some explicit dual controllers, such as the bicriterial and prediction error dual controllers [8], where the joint pdf $p(\mathbf{s}_k, \boldsymbol{\theta}_k | \mathbf{I}_k)$ is approximated by employing partial certainty equivalence [2]. Another possibility is to use approximation of the expectation value calculation. In this subsection an approximation based on the stochastic integration rule (SIR) [11], [13] will be presented.

The SIR evaluates the integral of the following form

$$\boldsymbol{\mu} = \int_{\mathbb{R}^{n_x}} \boldsymbol{\varphi}(\mathbf{x}) (2\pi)^{-n_x/2} \mathbf{e}^{-\frac{1}{2}\mathbf{x}^T \mathbf{x}} d\mathbf{x} \quad (23)$$

where $\boldsymbol{\varphi}(\cdot)$ is an arbitrary function. This integral (23) can be interpreted as computation of the expectation of the function $\boldsymbol{\varphi}(\mathbf{x})$, where \mathbf{x} is a random variable with pdf $p(\mathbf{x}) = \mathcal{N}(\mathbf{0}, \mathbf{I})$. However it can be trivially transformed into the form

$$\boldsymbol{\mu} = \int_{\mathbb{R}^{n_x}} \bar{\boldsymbol{\varphi}}(\mathbf{x}) \mathcal{N}(\hat{\mathbf{x}}, \mathbf{S}_x^T \mathbf{S}_x) d\mathbf{x} \quad (24)$$

where

$$\bar{\boldsymbol{\varphi}}(\mathbf{x}) = \boldsymbol{\varphi}(\mathbf{S}_x \mathbf{x} + \hat{\mathbf{x}}) \quad (25)$$

which has the following meaning $\boldsymbol{\mu} = E\{\bar{\boldsymbol{\varphi}}(\mathbf{x})\}$.

The stochastic integration rule proposed for solution of (23) 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 $N = 0$, initial value of the integral $\tilde{\boldsymbol{\mu}} = \mathbf{0}$, and compute the point $\chi_0 = \boldsymbol{\varphi}(\hat{\mathbf{x}})$

Step 3: Repeat (until $N = N_{max}$) the following loop:

- a) Set $N = N + 1$.
- b) Generate a uniformly random orthogonal matrix $\mathbf{U} \in \mathbb{R}^{n_x \times n_x}$ and generate a random number ρ from the chi distribution, i.e. $\rho \sim Chi(n_x + 2)$.
- c) Compute a set of points χ_i according to

$$\chi_i = -\rho \mathbf{S}_x \mathbf{U} \mathbf{e}_i + \hat{\mathbf{x}} \quad (26)$$

$$\chi_{n_x+i} = \rho \mathbf{S}_x \mathbf{U} \mathbf{e}_i + \hat{\mathbf{x}} \quad (27)$$

$$i = 1, \dots, n_x$$

and corresponding weights ω_i as

$$\omega_0 = 1 \frac{n_x}{\rho^2}, \quad (28)$$

$$\omega_i = \omega_{n_x+i} = \frac{1}{2\rho^2}. \quad (29)$$

- d) Compute the value \mathcal{J} of the integral at current iteration

$$\mathcal{J} = \sum_{i=0}^{2n_x} \boldsymbol{\varphi}(\chi_i) \omega_i \quad (30)$$

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

$$\tilde{\boldsymbol{\mu}} = \tilde{\boldsymbol{\mu}} + (\mathcal{J} - \tilde{\boldsymbol{\mu}}) / N. \quad (31)$$

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

The SIR can be understood as an improved version of the perfect Monte Carlo (MC) method. The SIR provides a better performance with respect to the MC especially in situations where the function $\boldsymbol{\varphi}(\cdot)$ is not approximately constant [11] or where the MC may have low accuracy and slow convergence. It is necessary to note that the SIR is based on randomly generated points and thus its result is also random but with an important property that the result is asymptotically exact, i.e. $E\{\tilde{\boldsymbol{\mu}}\} = \boldsymbol{\mu}$.

It should also be noted that for deterministically chosen orthogonal matrix $\mathbf{U} = \mathbf{I}$ and fixed $\rho = \sqrt{n_x + \kappa}$ the SIR matches the unscented transform (UT) [12] with κ representing the scaling parameter of the UT. However, the approximation based on UT provides only an approximate value of the expectation. The advantage of the SIR lies in the fact that the results are asymptotically exact.

C. Implicit control algorithm

The solution to the optimization recursion suggested in Subsection IV-A accompanied by the SIR method for evaluation on the expectations (14)-(16) constitutes the foundations of the new dual control algorithm proposed in this subsection.

The aim of the algorithm is to describe the technique for calculation of the dual control law defined by (20). Unfortunately, a closed form formula cannot be derived, that would define the control minimizing the expectation $E\{\mathcal{L}_k(\mathbf{s}_k, \boldsymbol{\theta}_k, \mathbf{u}_k) + \mathcal{V}_1 | \mathbf{u}_k, \mathcal{J}_k\}$. Thus, it is unavoidable to seek the minimizing control \mathbf{u}_k using a numerical optimization method. The following simple algorithm describes the procedure of finding the dual control at time instant k :

Algorithm 2: Dual control algorithm

Step 1: Determine the mean $E\{\mathbf{s}_k, \boldsymbol{\theta}_k | \mathcal{J}_k\}$ and the covariance matrix $\text{cov}\{\mathbf{s}_k, \boldsymbol{\theta}_k | \mathcal{J}_k\}$ of the filtering state estimate.

Step 2: Set the iteration counter $\ell = 0$.

Step 3: Choose an initial candidate for the suboptimal control $\mathbf{u}_k^{(0)}$.

Step 4: Repeat the following loop until a satisfactory minimizing control $\mathbf{u}_k^{(\ell)}$ is found

- a) Determine the mean $E\{\mathbf{s}_{k+1}^{(\ell)}, \boldsymbol{\theta}_{k+1}^{(\ell)} | \mathbf{u}_k^{(\ell)}, \mathcal{J}_k\}$ and the covariance matrix $\text{cov}\{\mathbf{s}_{k+1}^{(\ell)}, \boldsymbol{\theta}_{k+1}^{(\ell)} | \mathbf{u}_k^{(\ell)}, \mathcal{J}_k\}$ of the predictive state estimate.
- b) Calculate the approximation of Bellman function $\mathcal{V}_1^{(\ell)}$ using the relation (22).
- c) Evaluate the cost-to-go of the modified optimization problem stated in Subsection IV-A

$$J_k^{(\ell)}(\cdot) = E\left\{\mathcal{L}_k^{(\ell)}(\mathbf{s}_k, \boldsymbol{\theta}_k, \mathbf{u}_k^{(\ell)}) + \mathcal{V}_1^{(\ell)} | \mathbf{u}_k^{(\ell)}, \mathcal{J}_k\right\}.$$

- d) Test whether the relative change of the cost-to-go $J_k^{(\ell)}(\cdot)$ is satisfactory low, i.e., $J_k^{(\ell)}(\cdot) \approx \mathcal{V}_0$. If it is not, increment the counter ℓ , choose a new control candidate $\mathbf{u}_k^{(\ell)}$ and proceed with Step 4a).

The means and covariance matrices occurring in Steps 1 and 4 can be determined by employing any suitable estimation technique. However, because of the use of the SIR for the estimation of the expected values which was necessary for the cost-to-go evaluation, the use of the Stochastic integration filter [12], that is also based on the SIR, would be advisable. The benefit of this choice besides good quality of the estimates would be in the possibility to reuse the points generated by SIR.

The initial candidate for the minimizing control $\mathbf{u}_k^{(0)}$ in Step 3 can be determined using for example the cautious or the HCE controller. The SIR approximation is employed in Steps 4b) and 4c) where it is necessary to evaluate the expectations (14)-(16). The minimization of the cost-to-go in Step 4 is accomplished using a suitable numerical optimization method.

V. NUMERICAL EXAMPLE

This section is devoted to a numerical comparison of the new implicit dual controller (IDC) performance. The proposed controller will be compared with two basic passive (non-dual) controllers namely the often used HCE controller and the partial certainty equivalence (PCE) controller [2] which is of the cautious type and with two explicit dual controllers. The dual controllers are the bicriterial dual controller (BC) [7], [8] and the prediction error dual controller (PEDC) [8].

The performance of the individual controllers will be compared based on the estimate of the criterion value and on the root mean square error (RMSE) between the state and the state setpoint. The estimate of the criterion value is defined by relation

$$\hat{J} = \frac{1}{M} \sum_{j=1}^M J_j, \quad (32)$$

where J_j denotes the sum of the cost functions for particular system trajectory in one Monte Carlo run, i.e.

$$J_j = \sum_{k=0}^{N-1} \mathcal{L}_i(s_i, \theta_i, u_i). \quad (33)$$

The RMSE is evaluated using

$$\text{RMSE} = \frac{1}{M} \sum_{j=1}^M \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (s_{i+1} - \bar{s}_{i+1})^2} \quad (34)$$

The number of performed Monte Carlo runs is set to $M = 200$. It was assured that every controller operated on system driven by identical sequence of the state and measurement noises.

As a measure of the quality of the estimate \hat{J} , the variance of the criterion value estimate $\text{var}\{\hat{J}\}$ is presented. The value of $\text{var}\{\hat{J}\}$ is determined using the bootstrap technique [14].

The goal of the optimization problem is to find a control law minimizing the criterion

$$J = E \left\{ \sum_{k=0}^{N-1} (s_{k+1,2} - 5)^2 + 0.001 \cdot u_k^2 \right\}, \quad (35)$$

with respect to the discrete time stochastic system

$$s_{k+1} = \begin{pmatrix} 0 & 1 \\ \theta_{1k} & \theta_{2k} \end{pmatrix} s_k + \begin{pmatrix} 0 \\ \theta_{3k} \end{pmatrix} u_k + \mathbf{w}_k, \quad (36)$$

$$y_k = (0 \ 1) s_k + v_k. \quad (37)$$

where the unobservable parameters $\boldsymbol{\theta}_k = (\theta_{1,k}, \theta_{2,k}, \theta_{3,k})^T$ are considered as constant, i.e. $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k$. The actual parameters are given as $\boldsymbol{\theta}_k = (-2.0427, 0.3427, 1)^T$, $\forall k$ (i.e. the controlled system is unstable) and the initial state is fixed to value $s_0 = (1, -0.5)^T$.

The pdf's of the state and measurement noises are specified as

$$\mathbf{w}_k \sim \mathcal{N}(0, 0.0001), \quad (38)$$

$$v_k \sim \mathcal{N}(0, 0.001). \quad (39)$$

It should be noted that the solution of this optimization problem cannot be found in closed form.

The state and parameter estimates are delivered by the square root EKF where the initial condition $\hat{\mathbf{x}}_0$ for filter is considered to be a Gaussian random variable according to

$$\hat{\mathbf{x}}_0 \sim \mathcal{N}\left((1, -0.5, -2.0427, 0.3427, 1)^T, P_0\right) \quad (40)$$

with the initial covariance matrix $P_0 = 0.2 \cdot \mathbf{I}_6$.

All the compared suboptimal control laws will be determined employing the receding horizon cost-to-go criterion (9) with the lengths of the receding horizon set to $m = 2$ and the length of the examined horizon is $N = 20$.

The passive controllers will use the cost function

$$\mathcal{L}_i(s_i, \boldsymbol{\theta}_i, u_i) = (s_{i+1,2} - 5)^2 + 0.001 \cdot u_i^2 \quad (41)$$

i.e. it is essentially the criterion (5) with cost function (7) where

$$\mathcal{Q}_{k+1} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (42)$$

and $\mathbf{R}_k = 0.001$.

The bicriterial controller minimizes the criterion

$$J_k^e(\mathbf{u}_k) = -E \left\{ (s_{k+1} - \hat{s}_{k+1|k})^T \mathbf{W}_{k+1} (s_{k+1} - \hat{s}_{k+1|k}) \middle| \mathcal{J}_k \right\}, \quad (43)$$

on interval around the cautious control determined using the PCE control law given as

$$\boldsymbol{\Omega}_k = [u_k^C - \eta \text{tr} \mathbf{P}_k^{\theta\theta}, u_k^C + \eta \text{tr} \mathbf{P}_k^{\theta\theta}], \quad (44)$$

where \mathbf{W}_{k+1} is identity matrix, u_k^C represents the PCE control and design parameter is $\eta = 0.68$.

The PEDC considers the cost function as

$$\mathcal{L}_i(s_i, \boldsymbol{\theta}_i, u_i) = (s_{i+1,2} - 5)^2 + 0.001 \cdot u_i^2 + \bar{\mathbf{v}}_{i+1}^T \bar{\mathbf{L}}_{i+1} \bar{\mathbf{v}}_{i+1} \quad (45)$$

TABLE I
THE COMPARISON OF QUALITY OF THE COMPARED CONTROLLERS

Controller	RMSE	\hat{J}	$\sigma(J_i)$	$\text{var}\{\hat{J}\}$	Improv. %
LQG	1.24767	1.7221	0.0618	0.00002	31%
HCE	1.46623	2.4963	2.3732	0.0253	-
PCE	1.39205	2.1010	0.2498	0.0003	15%
BC	1.38114	2.0711	0.2431	0.0003	17%
PEDC	1.37707	2.0802	0.4329	0.0009	17%
IDC	1.36343	2.0329	0.2916	0.0004	19%

with \bar{L}_{i+1} defined as

$$\bar{L}_{i+1} = \bar{Q}_{i+1} - \bar{\Lambda}_{i+1}. \quad (46)$$

The matrices \bar{Q}_{i+1} and $\bar{\Lambda}_{i+1}$ were chosen as follows

$$\bar{Q}_{i+1} = \begin{pmatrix} \mathcal{Q}_{i+1} & \boldsymbol{\theta} \\ \boldsymbol{\theta} & \boldsymbol{\theta} \end{pmatrix}, \quad \bar{\Lambda}_{i+1} = \begin{pmatrix} 0.74\mathcal{Q}_{i+1} & \bar{\Lambda}_{i+1}^{s,\theta} \\ (\bar{\Lambda}_{i+1}^{s,\theta})^T & \boldsymbol{\theta} \end{pmatrix}, \quad (47)$$

where

$$\bar{\Lambda}_{i+1}^{s,\theta} = \begin{pmatrix} 0 & 0 & 0 \\ -0.3 & -0.3 & -0.4 \end{pmatrix}. \quad (48)$$

The zero elements of the matrix $\bar{\Lambda}_{i+1}^{s,\theta}$ in the final control law weigh the cross covariances between the parameters; the non-zero elements weigh the parameter autocovariances. The value of these non-zero elements influences the magnitude of probing.

The results of the Monte Carlo runs are summarized in Table I. The table compares the value of the RMSE, the criterion estimate \hat{J} , the variance of this estimate $\text{var}\{\hat{J}\}$, the standard deviation of cost of particular trajectories $\sigma\{J_i\}$ and improvement in percents compared to the HCE controller. Besides the above mentioned controllers the table also shows results for the unattainable LQG controller (i.e. for case when the parameters $\boldsymbol{\theta}_k$ are known).

It can be clearly seen that all the dual controllers provide better overall control quality compared to the passive non-dual controllers. This is due to the fact that the supplemented probing improves the quality of the parameter estimates. The proposed implicit dual controller provides the best performance among the compared dual controller. Moreover, the advantage of this controller is that it is not necessary to tune the design parameters which can often be a not straightforward and intuitive process. As a disadvantage an increase in computational demands can be seen due to the use of two numerical procedures, the stochastic integration rule and numerical minimization.

VI. CONCLUSION

The paper proposed the design of a new implicit dual controller. This dual controller provides a suboptimal solution to the optimization problem stated in the paper. The problem of closed form solvability of the BOR and of determining the expectation was discussed. The first problem was tackled by reformulation of the problem to the receding horizon type of the optimization problem and by reduction to a two-stage

optimization. The answer to the second problem was found by employing the stochastic integration rule for approximate evaluation of expectations. Finally, a numerical algorithm for calculation of the control action was introduced.

The numerical example then showed the performance of the proposed implicit dual controller compared to both passive and explicit dual controllers. It was illustrated that the new controller provides the best performance and, compared to the explicit dual controllers, without the need of tuning additional design parameters. The only notable obstacle can be seen in the fact, that this controller has higher computational demands due to the use of the numerical approximate methods.

Further research will be focused on a possible reduction of the computational demands and on generalization of the algorithm so that it could be used for multiple stage optimization problems, i.e. for longer receding horizons.

REFERENCES

- [1] A. Feldbaum, *Optimal Control Systems*. New York: Academic Press, 1965.
- [2] N. Filatov and H. Unberhauen, "Adaptive predictive control policy for nonlinear stochastic systems," *IEEE Trans. Automat. Control*, vol. AC-40, no. 11, pp. 1943–1949, November 1995.
- [3] J. Shi and D. Apley, "A suboptimal N-step-ahead cautious controller for adaptive control applications," *Journal of Dynamic Systems, Measurement, and Control*, vol. 120, no. 3, p. 419, 1998.
- [4] J. Wieslander and B. Wittenmark, "An approach to adaptive control using real time identification," *Automatica*, vol. 7, pp. 211–217, 1971.
- [5] N. Filatov and H. Unbehauen, "Survey of adaptive dual control methods," *IEE Proc.-Control Theory Appl.*, vol. 147, no. 1, pp. 118–128, 2000.
- [6] R. Milito, C. Padilla, R. Padilla, and D. Cadorin, "An innovations approach to dual control," *IEEE Trans. Automat. Contr.*, vol. AC-27, no. 1, pp. 132–137, February 1982.
- [7] N. M. Filatov, H. Unbehauen, and U. Keuchel, "Dual pole-placement controller with direct adaptation," *Automatica*, vol. 33, no. 1, pp. 113–117, 1997.
- [8] M. Flidr and M. Šimandl, "Dual adaptive controllers based on partial certainty equivalence," in *Proceedings of the 18th IFAC World Congress*, IFAC. Milano, Italy: Elsevier, August 2011, p. 34573462.
- [9] E. Tse, Y. Bar-Shalom, and L. Meier, "Wide-sense adaptive dual control for nonlinear stochastic systems," *IEEE Trans. Automat. Control*, vol. AC-18, no. 2, pp. 98–108, April 1973.
- [10] D. Bayard and A. Schumitzky, "Implicit dual control based on particle filtering," *Int. J. Adapt. Control Signal Process.*, 2008.
- [11] A. Genz and J. Monahan, "Stochastic integration rules for infinite regions," *SIAM J. Sci. Comput.*, vol. 19, no. 2, p. 426439, 1998.
- [12] J. Dunik, O. Straka, and Šimandl, "Stochastic integration filter," *To appear in IEEE Transactions on Automatic Control*.
- [13] A. Genz and J. Monahan, "A stochastic algorithm for high-dimensional integrals over unbounded regions with Gaussian weight," *Journal of computational and applied mathematics*, vol. 112, no. 1-2, pp. 71–81, 1999.
- [14] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. Chapman and Hall, 1994.