Proceedings of Machine Learning Research vol 120:1–10, 2020 2nd Annual Conference on Learning for Dynamics and Control

Dual Stochastic MPC for Systems with Parametric and Structural Uncertainty

Elena Arcari[1] EARCARI@ETHZ.CH
Lukas Hewing[1] LHEWING@ETHZ.CH
Max Schlichting[2] MAXSCHLICHTING@WEB.DE
Melanie N. Zeilinger[1] MZEILINGER@ETHZ.CH
[1] *Institute for Dynamic Systems and Control, ETH Zurich, Zurich, Switzerland*
[2] *Department of Neurorobotics, University of Freiburg, Freiburg, Germany*

Editors: A. Bayen, A. Jadbabaie, G. J. Pappas, P. Parrilo, B. Recht, C. Tomlin, M. Zeilinger

Abstract

Designing controllers for systems affected by model uncertainty can prove to be a challenge, especially when seeking the optimal compromise between the conflicting goals of identification and control. This trade-off is explicitly taken into account in the dual control problem, for which the exact solution is provided by stochastic dynamic programming. Due to its computational intractability, we propose a sampling-based approximation for systems affected by both parametric and structural model uncertainty. The approach proposed in this paper separates the prediction horizon in a dual and an exploitation part. The dual part is formulated as a scenario tree that actively discriminates among a set of potential models while learning unknown parameters. In the exploitation part, achieved information is fixed for each scenario, and open-loop control sequences are computed for the remainder of the horizon. As a result, we solve one optimization problem over a collection of control sequences for the entire horizon, explicitly considering the knowledge gained in each scenario, leading to a dual model predictive control formulation.

Keywords: dual control, stochastic adaptive control, nonlinear predictive control

1. Introduction

Model-based control techniques are commonly used to enable complex systems to perform challenging tasks, however, their performance strongly depends on the model accuracy. Adaptive control methods address the problem of simultaneous learning and control, avoiding the need for costly offline (re-)identification. Classic adaptive control is typically based on a single model with *parametric* uncertainty (Åström and Wittenmark, 2008; Sastry et al., 1990). This description can be insufficient when the effect of the uncertainty is large, and modifies the overall structure of the model. For such *structural* uncertainty (Murray-Smith and Johansen, 1997; Narendra and Han, 2011), it is beneficial to highlight several operating modes that, for instance, distinguish nominal behavior from fault conditions. In multiple-model approaches, a controller is designed to switch online between a finite number of models representing distinct system behaviors in order to ensure adequate control under each possible scenario. Available techniques, however, typically estimate the current operating mode of the system, and switch between pre-defined model-based controllers, but do not enable *active* learning about structural uncertainty, or consider the trade-off between sev-

© 2020 E. Arcari, L. Hewing, M. Schlichting & M.N. Zeilinger.

eral likely models. For a comprehensive review of multiple-model control we refer to (Heirung et al., 2019).

Active adaptive control schemes explicitly take into account the probing effect that the control input has on the system, and therefore the knowledge to be gained over structural or parametric uncertainty. Excitation needed to learn the model is optimally traded-off with the control performance when applying dual control (Feldbaum, 1961), for which the exact solution is provided by stochastic dynamic programming (DP) (Bertsekas, 2017). Due to its computational complexity, developed techniques revert to approximations, which can be classified into two categories: explicit and implicit approaches. Explicit techniques (Marafioti et al., 2014; Heirung et al., 2017) generally introduce a heuristic term in the cost function that drives the input to explore. These methods are usually computationally efficient, but require extensive expert knowledge and explore heuristically, e.g. often the entire system is excited for identification regardless of the benefit for the specific task. Implicit techniques (Tse and Bar-Shalom, 1973; Bayard and Eslami, 1985; Klenske and Hennig, 2016), on the other hand, are based on analytic reformulations that give a theoretically sound approximation of the original dual control problem, but are usually computationally expensive. Since implicit methods are developed in a generic framework, they can in principle be applied to a large class of systems, while ensuring the dual effect as an intrinsic property. We refer to (Mesbah, 2018) for a more detailed overview of dual control.

In this paper, we develop an implicit dual control scheme, by using a sampling method based on approximate dynamic programming (ADP), and making use of a rollout approach (Bertsekas, 2017). Samples are used to generate a scenario tree associated with an optimization problem that provides optimal control sequences for each scenario, taking the gained knowledge explicitly into account. This formulation gives rise to a dual stochastic model predictive control problem that builds on ideas included in (Arcari et al., 2019b) and (Hanssen and Foss, 2015), in which similar schemes are applied to systems linear in the uncertain parameters. In this paper, the formulation is extended to the important case of systems with uncertainty in both the parameters and in the model structure. The performance of the approach is shown in a simulation example of an aircraft and compared to model predictive control (MPC) with passive adaptation.

2. Preliminaries

2.1. Problem Formulation

We consider nonlinear dynamic systems, which are characterized by a finite set of operating modes $\mathcal{M} = \{M^1, \dots, M^{n_m}\}$, e.g. corresponding to different failure cases. Each operating mode $M \in \mathcal{M}$ is associated with a prior probability $p(M)$ of being active, and the system dynamics given mode M are

$$M: \; x_{k+1} = \Phi^M(x_k, u_k)\gamma^M + w_k^M, \tag{1}$$

with state $x_k \in \mathbb{R}^{n_x}$ and control input $u_k \in \mathcal{U} \subseteq \mathbb{R}^{n_u}$, where $\mathcal{U}$ defines the input constraints. Each mode is described by a nonlinear basis function matrix $\Phi^M : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x \times n_\gamma^M}$ with associated uncertain parameters $\gamma^M \in \mathbb{R}^{n_\gamma^M}$, which we assume to be distributed according to $\gamma^M \sim \mathcal{N}(\mu_\gamma^M, \Sigma_\gamma^M)$. The system is furthermore subject to additive disturbances $w_k^M \sim \mathcal{N}(0, \Sigma_w^M)$, which we consider zero mean i.i.d. for (notational) simplicity. At each time step k, we assume access to measurements of the state x_k, which are used to identify both the uncertain operating mode M and the corresponding parameterization γ^M.

Remark 1 *The structural uncertainty description through a collection of modes $\mathcal{M}$ is applicable to a wide range of use cases. This includes, for instance, systems affected by parameters taking discrete values, multimodal parameter distributions or discretizations of continuous uncertain parameters which do not allow for an expression in the form of γ in (1).*

The considered objective is to find an optimal policy that simultaneously controls the system and identifies its operating mode and corresponding unknown parameters. For this purpose, we define a finite-horizon cost

$$J_{\bar{N}}(\Pi, x_0) := \mathop{\mathbb{E}}_{\substack{M,\gamma^M,\\ w_0^M,\ldots,w_{\bar{N}-1}^M}} \sum_{k=0}^{\bar{N}} l_k(x_k, \pi_k(x_k)), \tag{2}$$

where $\Pi = \{\pi_0(\cdot), \ldots, \pi_{\bar{N}-1}(\cdot)\}$ is a policy sequence that we want to optimize by minimizing (2), $\bar{N}$ is the length of the overall control task and $l_k : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}$ is a potentially time-varying cost function. As a result, we obtain a finite-horizon stochastic optimal control problem that can be addressed using dynamic programming (DP), which offers the property of generating dual control policies. In the following section we provide a DP formulation that handles both parametric and structural uncertainty, and subsequently outline a tractable approximate DP scheme.

2.2. Stochastic DP with Parametric and Structural Model Uncertainty

In order to address the problem by DP, we provide a formulation in the context of systems with imperfect state information (Bertsekas, 2017). We define the information vector $\mathcal{I}_k$, expressing the available information at time step k recursively given a known initial state , i.e. $\mathcal{I}_0 = x_0$, as

$$\mathcal{I}_{k+1} = \left[x_{k+1}^T, u_k^T, \mathcal{I}_k^T\right]^T.$$

The information at time step k takes the function of the state in the DP recursion

$$J_k^*(\mathcal{I}_k) = \min_{\pi_k} \; l_k(x_k, \pi_k) + \mathop{\mathbb{E}}_{M,\gamma^M,w_k^M} \left[J_{k+1}^*(\mathcal{I}_{k+1}) \mid \mathcal{I}_k\right], \quad k = 0, \ldots, \bar{N}-1 \tag{3}$$

initialized at $J_{\bar{N}}^*(\mathcal{I}_{\bar{N}}) = l_{\bar{N}}(x_{\bar{N}})$. The expected value arising in the recursion is computed with respect to the currently available information $\mathcal{I}_k$, which means that the parameter distribution is accordingly updated at each time step. When receiving a new state measurement x_{k+1}, which updates the current information vector to $\mathcal{I}_{k+1}$, the posterior joint distribution of mode M and γ^M is given by:

$$p(\gamma^M, M \mid \mathcal{I}_{k+1}) = p(\gamma^M \mid \mathcal{I}_{k+1}, M) p(M \mid \mathcal{I}_{k+1}),$$

where $p(M \mid \mathcal{I}_{k+1})$ is the probability that mode M is active given the information at time step $k+1$, and $p(\gamma^M | \mathcal{I}_{k+1}, M)$ is the distribution of the associated parameters. Each component is updated via Bayesian estimation (Mesbah, 2018):

$$p(\gamma^M \mid \mathcal{I}_{k+1}, M) = \frac{p(x_{k+1} \mid u_k, \mathcal{I}_k, M, \gamma^M) p(\gamma^M \mid \mathcal{I}_k, M)}{p(x_{k+1} \mid u_k, \mathcal{I}_k, M)}, \tag{4a}$$

$$p(M \mid \mathcal{I}_{k+1}) = \frac{p(x_{k+1} \mid u_k, \mathcal{I}_k, M) p(M \mid \mathcal{I}_k)}{p(x_{k+1} \mid u_k, \mathcal{I}_k)}. \tag{4b}$$

The parameter update given mode M in (4a) follows Gaussian linear regression, in which the likelihood $p(x_{k+1} \mid u_k, \mathcal{I}_k, M, \gamma^M)$ can be computed from the corresponding Gaussian noise distribution in (1). The normalization factor $p(x_{k+1} \mid u_k, \mathcal{I}_k, M)$ is subsequently used to evaluate the likelihood of each individual mode in (4b). For details on the Bayesian estimation procedure we refer to Appendix A in Arcari et al. (2019a).

Remark 2 *For long run times $\bar{N}$, the formulation above leads to convergence of both the operating mode and parameter estimate to absolute certainty, due to the accumulated information. This results in a loss of adaptability, which can be undesired e.g. for fault detection. Convergence can be avoided, for instance, by introducing process noise on the parameter γ^M and switching between modes according to a Markov process, resulting in increased complexity of the estimation in* (4)*, see e.g. (Maybeck, 1979). As a computationally efficient alternative, we heuristically cap the probability of each mode, as well as the variance of the parameter estimate.*

While DP provides the exact solution to the dual control problem in the case of both parametric and structural model uncertainty, it is generally tractable only for very small problems (Klenske and Hennig, 2016). We provide a tractable approximate approach based on (Arcari et al., 2019b), which we extend to the case of both continuous parameters and structural uncertainty.

3. Dual Stochastic MPC

The proposed formulation is based on a receding horizon implementation of an approximate dynamic programming (ADP) strategy. We repeatedly solve an approximation of problem (2) over a shortened horizon $N \ll \bar{N}$, and apply the first computed control input, i.e. using receding horizon control. To approximate the receding horizon problem, we furthermore use a rollout approach (Bertsekas, 2017), and truncate the prediction horizon of length N to $L < N$. The control sequence associated with the truncated part of the horizon of length $N - L$ is optimized in open-loop, as in model predictive control, and defines a suboptimal cost-to-go $\tilde{J}_L(\mathcal{I}_L)$. The first part of the horizon of length L, is solved by approximating the DP recursion (3) using a sampling-based approach. The expected values arising at each DP iteration in (3) are generally not available in closed form, even if the Bayesian updates in (4) can be evaluated exactly, e.g. with self-conjugate distributions (Klenske and Hennig, 2016). For this reason, in Arcari et al. (2019b), expected values are evaluated as averages over parameter and noise samples. We use a similar idea for handling the case of both continuous parameters and structural model uncertainty, such that it preserves the dual properties of the DP policy. In the following subsections, we outline the sampling approach used for approximating DP along the first L steps of the prediction horizon, which we refer to as *dual* part, and we discuss the cost-to-go $\tilde{J}_L(\mathcal{I}_L)$ that is solved in the *exploitation* part, covering the remaining part of the horizon of length $N - L$.

3.1. Dual Part

The dual part approximates (3) by computing the expected values as averages over samples of the noise w and of the parameter γ, weighted by the probability of each mode M. This induces an associated scenario tree depicted in Figure 1, in which each state node $x_k^{j_k}$, denoted by a black circle, generates together with its corresponding input $u_k^{j_k}$ an associated subtree, where $j_k = 1, \ldots, (N_s n_m)^k,\ k = 0, \ldots, L$. In each node, at every time step, there are two stages of

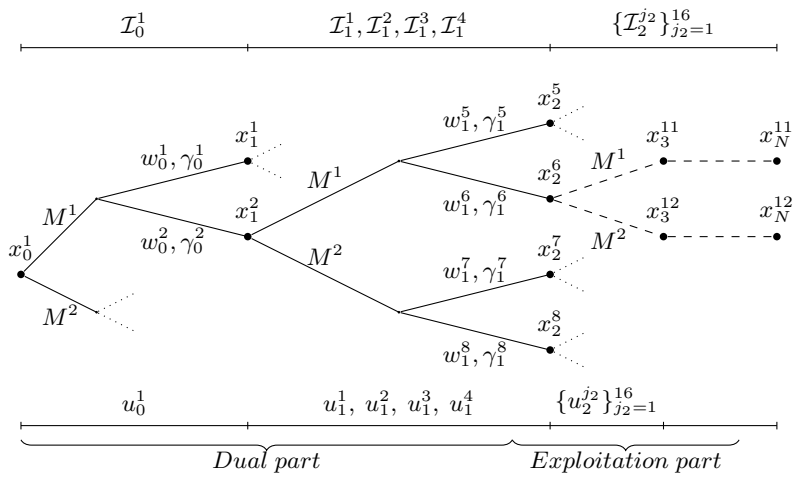

Figure 1: Scenario tree for $L = 2$: The dual part starts at a given initial condition and evolves, under a given input sequence, according to the $n_m = 2$ operating modes M_1, M_2, and the $N_s = 2$ noise w and parameter γ samples. The exploitation part further predicts the trajectory using the current parameter and mode distribution for steps $k = 2, \dots, N-1$.

branching. The first considers each potential mode M, while in the second, N_s noise and parameter samples are realized, generating in total $N_s \times n_m$ leaf nodes. This procedure is repeated iteratively along a horizon of length L, resulting in the following update for each state sample

$$x_{k+1}^{j_{k+1}} = \Phi^M(x_k^{P(j_{k+1})}, u_k^{P(j_{k+1})})\gamma_k^{j_{k+1}} + w_k^{j_{k+1}}, \quad j_{k+1} = 1, \dots, (N_s n_m)^{k+1} \tag{5}$$

where the index j_{k+1} refers to the samples at stage $k+1$, and $P(\cdot)$ indicates the parent function.

The sampled state realizations obtained from (5) are used to construct the information vector $\mathcal{I}_k^{j_k}$ associated with each node. The dual effect is obtained by using (4) to compute both the distribution of γ, i.e. $p(\gamma^M \,|\, \mathcal{I}_k^{j_k}, M)$, from which we generate the parameter samples, as well as the probability of each mode $p(M \,|\, \mathcal{I}_k^{j_k})$, which is used to weigh the cost when averaging over all nodes of a time step. The expected values in (3) are therefore substituted with weighted averages over these samples, which allows for unnesting the minimizations arising in the DP recursion, and optimizing over all the control input sequences in the tree at the same time, as was first proposed for handling parametric uncertainty in Arcari et al. (2019b). Therefore, while the DP recursion is not explicitly carried out, the input associated with each subtree is optimized while simultaneously considering future observations, and reacting to the corresponding sampled disturbance realizations, hence providing dual control. The optimization problem for the dual part is given by

$$\tilde{J}_0(I_0^1) := \min_{\boldsymbol{u}_0,\dots,\boldsymbol{u}_{L-1}} \bar{p}_{j_0} l_0(x_0^1, u_0^1) + \frac{1}{N_s}\sum_{j_1=1}^{N_s n_m} \bar{p}_{j_1} l_1(x_1^{j_1}, u_1^{j_1}) + \dots + \frac{1}{N_s^L}\sum_{j_L=1}^{(N_s n_m)^L} \bar{p}_{j_L} \tilde{J}_L(\mathcal{I}_L^{j_L}), \tag{6}$$

where $\boldsymbol{u}_k = \{u_k^1, \dots, u_k^{(N_s n_m)^k}\}$ refers to the collection of inputs for each stage of the problem, and $u_k^{j_k} \in \mathcal{U}$. The weight $\bar{p}_{j_k}$, assigned to each node, can be recursively obtained using the mode probability

$$\bar{p}_{j_{k+1}} = p(M|I_k^{P(j_{k+1})})\bar{p}_{P(j_{k+1})}, \quad k = 0, \dots, L-1 \tag{7}$$

with $\bar{p}_{j_0} = 1$. Details about the derivation of (6) can be found in Appendix B in Arcari et al. (2019a).

3.2. Exploitation Part

In the exploitation part, we formulate the approximate cost-to-go $\tilde{J}_L(\mathcal{I}_L^{j_L})$ for each scenario, by fixing the information collected until step L. Therefore, the current probability of each operating mode and its associated parameter distribution remain constant along each scenario in the exploitation part. Optimizing over these branches corresponds to solving a non-dual stochastic MPC problem

$$\begin{aligned}
&\tilde{J}_L(\mathcal{I}_L^{j_L}) = \min_{u_L^{j_L},\dots,u_{N-1}^{j_L}} \tilde{J}_L(\mathcal{I}_L^{j_L}, u_{L:N-1}^{j_L}),\\
&\tilde{J}_L(\mathcal{I}_L^{j_L}, u_{L:N-1}^{j_L}) = \sum_{m=1}^{n_m} p(M^m|\mathcal{I}_L^{P(j_{L+1})}) \Big(\mathbb{E}_{\gamma^M|M,w_L^M,\dots,w_{N-1}^M}\Big[l_L(x_L^{P(j_{L+1})}, u_L^{P(j_{L+1})}) +\\
&+ \sum_{k=L+1}^{N-1} l_k(x_k^{j_{L+1}}, u_k^{j_{L+1}}) + l_N(x_N^{j_{L+1}}) \mid \mathcal{I}_L^{P(j_{L+1})}\Big]\Big)
\end{aligned} \tag{8}$$

where the expected value with respect to the current mode probability, given the information at parent node $P(j_{L+1})$ with $j_{L+1} = 1,\dots,n_m(N_s n_m)^L$, is formulated explicitly. When the cost function is chosen to be either quadratic or linear, there exist analytic reformulations of the expected value in (8) with respect to the first two moments of the predicted state. A common approach consists in considering only mean information, and is generally referred to as certainty equivalence control, for which we consider (1) evaluated at the state and parameter mean

$$\mu_{x_{L+1}}^{j_{L+1}} = \Phi^M(x_L^{P(j_{L+1})}, u_L^{P(j_{L+1})})\mu_{\gamma_L}^{P(j_{L+1})} \tag{9a}$$

$$\mu_{x_{k+1}}^{j_{L+1}} = \Phi^M(\mu_{x_k}^{j_{L+1}}, u_k^{j_{L+1}})\mu_{\gamma_L}^{P(j_{L+1})}, \quad k = L+1,\dots,N-1 \tag{9b}$$

where $\mu_{\gamma_L}^{P(j_{L+1})}$ is fixed at the value obtained at the last dual step L, and for parent node $P(j_{L+1})$. A derivation of the model propagation in terms of state mean and variance information, using a first-order Taylor approximation of the dynamics around the state mean, is provided in Appendix C in Arcari et al. (2019a).

3.3. Final Dual MPC Problem

The overall approximate dual control formulation provides a setup in which the two subproblems, namely the dual and exploitation part, can be merged into one optimization problem that optimizes over control sequences along the whole prediction horizon of length N, and which can be solved in a receding horizon fashion:

$$\begin{aligned}
\min_{\boldsymbol{u}_0,\dots,\boldsymbol{u}_{N-1}} \quad & \sum_{k=0}^{L-1} \frac{1}{(N_s)^k} \sum_{j_k=1}^{(N_s n_m)^k} \bar{p}_{j_k} l_k(x_k^{j_k}, u_k^{j_k}) + \frac{1}{(N_s)^L} \sum_{j_L=1}^{(N_s n_m)^L} \bar{p}_{j_L} \tilde{J}_L(\mathcal{I}_L^{j_L}, u_{L:N-1}^{j_L})\\
\text{s.t.} \quad & (5) && k = 0,\dots,L-1,\\
& \gamma_k^{j_k} \text{ drawn from } \mathrm{P}(\gamma^M|M,\mathcal{I}_k^{j_k}) && k = 0,\dots,L-1,\\
& (7) && k = 0,\dots,L-1,\\
& (9) && k = L,\dots,N-1,\\
& u_k^{j_k} \in \mathcal{U} && k = 0,\dots,N-1
\end{aligned} \tag{10}$$

Parameter samples $\gamma_k^{j_k}$ are generated online as affine transformations of samples from a standard normal distribution, using exact mean and covariance provided by (4a), and can be drawn before optimization, such that standard solvers for gradient-based optimization can be used.

Remark 3 *The formulation of problem (10) offers the possibility to use deferred dual steps, i.e. inserting dual steps after a number of exploitation steps. This is useful for instance when the system requires several time steps to achieve good excitation, such that identification later in the prediction can be significantly more informative.*

4. Simulation Example

The following example demonstrates the developed dual MPC (DMPC) for the problem of controlling an aircraft to a desired altitude, with a reference change of $50\ m$. The nominal operating mode of the system follows the continuous-time linearized longitudinal dynamics of the Cessna Citation aircraft in (Maciejowski, 2002). The state of the system x_k describes the evolution of the angle of attack, the pitch angle, the pitch rate and the altitude. The input u_k corresponds to the elevator angle, and is constrained to lie between $\pm 0.2\ rad$. As a simulation scenario, we consider the case in which a fault occurs to the actuator some time before the reference change, leading to a decrease in gain of 75%. We identify two operating modes of the system with M^1 being nominal and M^2 being the fault mode, each with a corresponding uncertain parameter γ^M describing the actuator gain

$$M:\ x_{k+1} = Ax_k + \gamma^M Bu_k + w_k, \quad M \in \{M^1, M^2\}$$

where A, B are the discretized system matrices using a sampling time $T_s = 0.2\ s$. We assume that γ^2 is subject to larger uncertainty than γ^1, as it is hard to know a priori how much loss in gain has occurred. For this reason we choose the prior distribution for each parameter as $\gamma^1 \sim \mathcal{N}(0.95, (0.03)^2)$ and $\gamma^2 \sim \mathcal{N}(0.4, (0.1)^2)$. Both modes are subject to additive process noise $w_k \sim \mathcal{N}(0, \Sigma_w)$, with $\Sigma_w = \text{diag}([(0.01)^2, (0.01)^2, (0.1)^2, (0.01)^2])$. Compared with the use of a unimodal parameter distribution with a single mode, explicitly considering a failure case provides faster model identification, as shown for instance in (Boskovic and Mehra, 1999). We generate a DMPC with dual lookahead $L = 1$, $n_m = 2$ operating modes, and $N_s = 2$ noise and parameter realizations, which result in $N_s \times n_m = 4$ scenarios, and the overall prediction horizon is $N = 20$. The cost function is quadratic, and the corresponding weights on state and input are chosen as $Q = \text{diag}([0, 0, 0, 1])$ and $R = 1000$ respectively. We compare the DMPC with a non-dual, passive adaptive controller, referred to as CEMPC. This is based on a certainty equivalence approach, i.e. a deterministic receding horizon controller that computes open-loop sequences, and in which the probability of the operating mode and the underlying distribution of the associated unknown parameter are updated in closed-loop following (4). Figure 2 shows the performance of both controllers for the specified task, where the first dashed line refers to the fault occurring at time step $k = 20$, and the second dashed-dotted line corresponds to the reference change at time step $k = 60$. For DMPC, the elevator input (top right) explores aggressively as soon as the reference change enters the prediction horizon at time step $k = 40$, and displays goal-oriented control, i.e. fast identification of a fault ensuring adequate performance by initiating the pitching maneuver early. On the other hand, the non-dual CEMPC input does not excite the system sufficiently in advance, causing input saturation and late response to the reference change. Suboptimality of this behavior is quantified in terms of the closed-loop cost averaged over 200 noise realizations as $J_{DMPC} = 52.5$ and

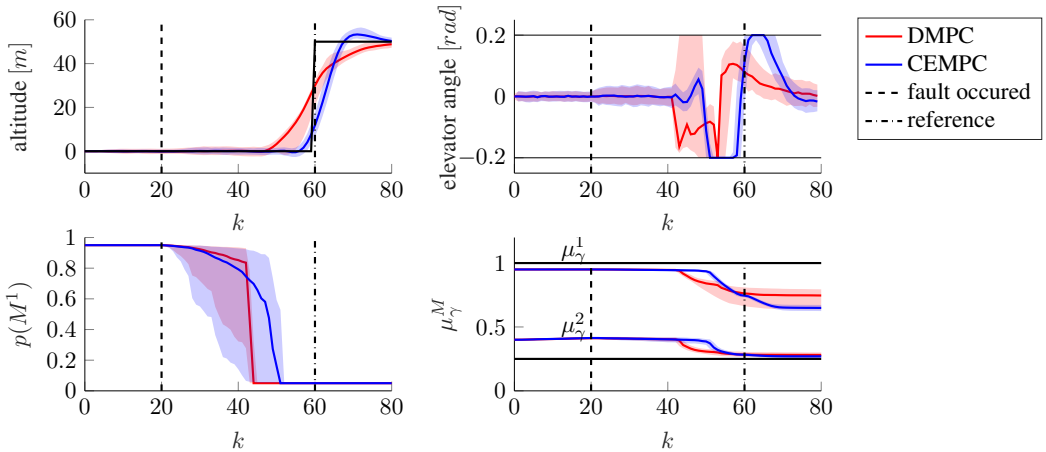

Figure 2: Altitude reference change. Each plot depicts the DMPC (in red) and CEMPC (in blue) performance averaged over 200 noise realizations, for which the median is highlighted in bold. The shaded area around the median is bounded by the 5 and 95 percent quantiles. *Top left*: altitude state trajectory. *Top right*: elevator input. *Bottom left*: nominal mode probability. *Bottom right*: parameter mean for both modes.

$J_{CEMPC} = 62.65$. The probability of the nominal mode (bottom left) in the DMPC shifts in one time step at $k = 40$ to $p(M^1) = 0.05$, which corresponds to the chosen lower bound (cf. Remark 2). The DMPC bimodal distribution defined by the two parameters γ^1 and γ^2 starts to adapt due to introduced excitement at $k = 40$ towards the true parameter $\gamma^2 = 0.25$, associated with the fault mode active at the reference change. Convergence of μ_γ^1 slows down as the aircraft approaches the reference since input excitation decreases, i.e. further identification is no longer needed. Similarly, the estimate for γ^2 stagnates due to little excitation, but would eventually converge to 0.25. The DMPC controller therefore provides rapid identification of the fault condition, and of the specific loss of the actuator gain.

5. Conclusion

We have presented an approximate dual MPC approach for active learning and control of systems affected by parametric and structural uncertainty. The formulation is based on a scenario tree approach, constructed by sampling the noise and the unknown parameters associated to each operating mode of the system. The aircraft control example illustrates faster identification of the active mode and of the unknown parameters, while ensuring a smooth maneuver towards the desired altitude.

Acknowledgments

This work was supported by the Swiss National Science Foundation under grant no. PP00P2 157601 / 1 and by the CTI-Project Nr. 25959.2 PFIW-IW.

References

Elena Arcari, Lukas Hewing, Max Schlichting, and Melanie N. Zeilinger. Dual stochastic MPC for systems with parametric and structural uncertainty. *arXiv preprint*, 2019a.

Elena Arcari, Lukas Hewing, and Melanie N. Zeilinger. An approximate dynamic programming approach for dual stochastic model predictive control. *arXiv1911.03728*, 2019b.

Karl J Åström and Björn Wittenmark. *Adaptive Control*. Dover Publications, 2 edition, 2008.

David S. Bayard and Mansour Eslami. Implicit dual control for general stochastic systems. *Optimal Control Applications and Methods*, 6(3):265–279, 1985.

Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 4th edition, 2017. ISBN 1886529264.

J. D. Boskovic and R. K. Mehra. Stable multiple model adaptive flight control for accommodation of a large class of control effector failures. In *Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251)*, volume 3, pages 1920–1924 vol.3, June 1999.

A. A. Feldbaum. Dual control theory. Part I & II. *Automation and Remote Control*, 21(9):874–880, 1033–1039, 1961.

Kristian G. Hanssen and Bjarne Foss. Scenario based implicit dual model predictive control. *IFAC-PapersOnLine*, 48(23):416 – 421, 2015. ISSN 2405-8963. 5th IFAC Conference on Nonlinear Model Predictive Control NMPC 2015.

Tor Aksel N. Heirung, B. Erik Ydstie, and Bjarne Foss. Dual adaptive model predictive control. *Automatica*, 80:340–348, 2017. ISSN 0005-1098.

Tor Aksel N. Heirung, Tito L.M. Santos, and Ali Mesbah. Model predictive control with active learning for stochastic systems with structural model uncertainty: Online model discrimination. *Computers & Chemical Engineering*, 128, 05 2019.

Edgar D. Klenske and Philipp Hennig. Dual control for approximate Bayesian reinforcement learning. *J. Machine Learning Research*, 17(1):4354–4383, January 2016. ISSN 1532-4435.

J.M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, 2002.

Giancarlo Marafioti, Robert R. Bitmead, and Morten Hovd. Persistently exciting model predictive control. *Int. J. Adaptive Control and Signal Processing*, 28(6):536–552, 2014.

P.S. Maybeck. *Stochastic Models, Estimation, and Control*. Mathematics in science and engineering. Academic Press, 1979. ISBN 9780124807020.

Ali Mesbah. Stochastic model predictive control with active uncertainty learning: A survey on dual control. *Annual Reviews in Control*, 45:107 – 117, 2018. ISSN 1367-5788.

Roderick Murray-Smith and T.A. Johansen. *Multiple model approaches to nonlinear modelling and control*. CRC press, 1997.

Kumpati S. Narendra and Zhuo Han. The changing face of adaptive control: the use of multiple models. *Annual reviews in control*, 35(1):1–12, 2011.

Shankar Sastry, Marc Bodson, and James F. Bartram. Adaptive control: Stability, convergence, and robustness. *The Journal of the Acoustical Society of America*, 88(1):588–589, 1990.

E. Tse and Y. Bar-Shalom. An actively adaptive control for linear systems with random parameters via the dual control approach. *Trans. Automatic Control*, 18(2):109–117, April 1973. ISSN 0018-9286.