SELF-REFLECTIVE MODEL PREDICTIVE CONTROL

BORIS HOUSKA[1,2] , DRIES TELEN[3] , FILIP LOGIST[3] , AND JAN VAN IMPE[3]

Abstract. This paper proposes a novel control scheme, named self-reflective model predictive control, which takes its own limitations in the presence of process noise and measurement errors into account. In contrast to existing output-feedback MPC and persistently exciting MPC controllers, the proposed self-reflective MPC controller does not only propagate a matrix-valued state forward in time in order to predict the variance of future state-estimates, but it also propagates a matrix-valued adjoint state backward in time. This adjoint state is used by the controller to compute and minimize a second order approximation of its own expected loss of control performance in the presence of random process noise and inexact state estimates. The properties of the proposed controller are illustrated with a small but non-trivial case study.

Key words. Optimal Control, Model Predictive Control, Optimal Experiment Design, Dual Control

AMS subject classifications. 49K21, 49K30, 93B07, 93B52

1. Introduction. Historically, online optimization based receding horizon controllers have been developed by separating the overall control problem into two, namely, a state estimation and a state feedback control problem [1, 49]. Nowadays, there exists a mature theoretical foundation as well as software for both moving horizon estimation (MHE) [12, 48], realizing optimized state estimation, as well as model predictive control (MPC) [6, 7, 14, 28], realizing optimized state feedback control.

Unfortunately, if the dynamic system is not linear or if there are state- or control constraints presents, a separation of the overall control problem into an estimation and a feedback phase can be sub-optimal. For example, if a nonlinear system is not observable at its desired operation point, it might be necessary to excite the system from time to time on purpose in order to be able to measure the states. One way to address this problem is by using a so-called output-feedback model predictive controller, as suggested in [18, 41, 42]. During the last decade, there has been significant progress in the field of output-feedback MPC. In particular, in [21] a way of approaching the output-feedback MPC problem for linear systems in the presence of convex state- and control constraints has been developed. The corresponding method uses convex optimization techniques in order to design a feedback law that is affine in the disturbance sequence by solving an online optimization problem whose complexity scales quadratically with respect to the prediction horizon. Other approaches for conservatively solving the output-feedback problem based on set propagation techniques can be found in [5, 11, 33, 36, 57].

Notice that the output-feedback model predictive control problem is related to the so-called *dual control* problem [15]. Dual control tries to find an optimal control input under the assumption that learning is possible while the model is unknown. A complete overview about the historical developments related to the dual control problem is beyond the scope of this article, but there exists a large number of recommendable overview articles [16, 17, 37, 55, 56], where reviews and comments about the history and importance of the dual control problem can be found. An early article

[1]Corresponding Author (borish@shanghaitech.edu.cn).

[2]School of Information Science and Technology (SIST), ShanghaiTech University, 319 Yueyang Road, Shanghai 200031, China.

[3]Chemical Engineering Department, KU Leuven, BioTeC+ & OPTEC, Gebroeders De Smetstraat, 9000 Ghent, Belgium.

about the limitations of predictive control from an adaptive control perspective can be found in [8]. During the last decade, a number of articles have appeared, which all try in one or the other way to include dual control aspects in MPC, as reviewed below. Here, one way to predict future state- and parameter estimation errors is to use an extended Kalman filter. This idea has been introduced by Hovd and Bitmead [31], who propose to augment standard MPC by an additional matrix-valued state that predicts the variance of future state estimates. This matrix-valued state is propagated forward in time by a Riccati differential equation, which computes the variance state of an extended Kalman filter [53]. It can be shown that the corresponding controller excites the system for improved state estimates. A related approach toward dual MPC has been proposed in [24, 25], where the MPC objective is augmented by a term that—similar to optimal experiment design [19, 46, 54]—penalizes the predicted parameter error variance. In [26] the same authors develop a way to predict and optimize future parameter estimation errors for single-input-single-output finite impulse response systems based on non-convex quadratically constrained quadratic programming formulations. A different way of taking predicted parameter estimation errors into account has been suggested in [32], where an MPC scheme is augmented by an adaptive parameter estimation algorithm. Yet another related control problem formulation is proposed in [34], where it is suggested to add an experiment design constraint to the MPC formulation. Lucia and Paulen [38] suggest to combine guaranteed parameter estimation and multi-stage MPC techniques in order to reduce the uncertainty about future estimation errors. Moreover, in [20] and [35] various extensions of application oriented experiment design are developed in order to generate excitation signals in the context of receding horizon control. Another recent article [43] on online experiment design suggests to use a modified experiment design framework for generating "least costly" excitations, which try to avoid perturbation of the nominal plant operation as much as possible. Other approaches for persistently exciting closed loop MPC can be found in [27, 39, 51, 58].

1.1. Contribution. After introducing the mathematical notation and background in Section 2, the paper starts with Section 3 by outlining an example, which illustrates why MPC may fail to stabilize nonlinear control systems if the MPC controller does not take the properties of the state estimator into account. Sections 4 and 5 review existing concepts from the field of MPC, dynamic programming, and state estimators based on extended Kalman filters and MHE. The main theoretical contribution of this paper is presented in Section 6, which proposes a novel way of analyzing the loss of optimality that is associated with implementing certainty equivalent MPC controllers in the presence of process disturbances and measurement noise. This analysis is related to a recently proposed method for formulating economic optimal experiment design problems [29]. However, [29] analyzes a static optimization problem while this paper is concerned with the loss of optimality of a dynamic two-player-multi-stage game, as the MPC controller is optimizing the control input whenever new state estimates become available. A major contribution in this context is presented in Theorem 6.3, which proposes a procedure to compute second order moment expansions of the expected loss of optimality of certainty equivalent MPC in the presence of random process noise as well as random measurement errors. The corresponding method scales linearly with the prediction horizon of the MPC controller despite the fact that the future state estimates are correlated in time. In Section 7 this second order moment approximation of the expected sub-optimality of certainty equivalent MPC is used to design a novel "self-reflective" model predic-

tive controller, which optimizes not only its nominal performance but also includes a term that approximates its own expected sub-optimality in the presence of process noise and measurement errors. This leads—similar to the above reviewed dual MPC schemes—to a controller that persistently excites the states in order to reduce the variance of future state estimates. However, the above reviewed existing dual MPC schemes are all based on additional matrix-valued hyperstates that propagate forward in time in order to predict the accuracy of future state estimates. The proposed receding horizon controller also propagates states forward in time, but, in contrast to existing dual MPC schemes, it is based on additional matrix-valued adjoint states, which propagate backward in time. By coupling the information of these forward and adjoint states the control scheme obtains the ability to be "self-reflective", in the sense that it is capable of taking its own limitations into account. The novelty and differences of this approach compared to the numerous existing articles on persistently exciting MPC is further discussed toward the end of Section 7. Section 8 presents a numerical case study, which illustrates the properties of self-reflective MPC. Section 9 concludes the paper.

1.2. Notation. For any matrix $M \in \mathbb{R}^{n \times n}$, the notation $M \succ 0$ indicates that M is positive definite. The notation $M^\dagger$ is used to denote the pseudo inverse of a matrix M. If $\varphi : \mathbb{R}^n \to \mathbb{R}^m$ is an integrable function and x a random variable with integrable probability distribution $\rho : X \to \mathbb{R}$ and support $X \subseteq \mathbb{R}^n$, the expected function value of φ is denoted by

$$\mathbb{E}_x\{\varphi(x)\} \;=\; \int_X \varphi(x)\rho(x)\,\mathrm{d}x \;.$$

Notice that the index is occasionally omitted, i.e., we simply write $\mathbb{E}\{\varphi(x)\}$, if it is clear from the context that x is the random variable with respect to which the expected value is computed. If the random variable has a block structure,

$$x = (x_1^\mathsf{T}, x_2^\mathsf{T})^\mathsf{T} \quad \text{with} \quad X_1(x_2) = \left\{ x_1 \in \mathbb{R}^{n_1} \mid (x_1^\mathsf{T}, x_2^\mathsf{T})^\mathsf{T} \in X \right\}$$

and $0 \leq n_1 \leq n$, the conditional expectation over x_1 is denoted by

$$\mathbb{E}_{x_1}\{\varphi(x)\} \;=\; \int_{X_1(x_2)} \varphi(x)\rho(x)\,\mathrm{d}x_1 \;.$$

Notice that $\mathbb{E}_{x_1}\{\varphi(x)\}$ is a function of x_2. In particular, this notation is such that

$$\mathbb{E}_{x_1}\mathbb{E}_{x_2}\{\varphi(x)\} = \mathbb{E}_{x_2}\mathbb{E}_{x_1}\{\varphi(x)\} = \mathbb{E}_x\{\varphi(x)\} \;.$$

In addition, the notation $\| \cdot \|_2$ denotes either the Euclidean norm, if the argument is a vector, or the spectral norm, if the argument is a quadratic matrix. Moreover, we use the upper index notation

$$x^{[k]} = (x_0, x_1, x_2, \ldots, x_k)$$

to denote the sub-matrix that consists of the first $k + 1$ columns of a given matrix $x = (x_0, x_1, \ldots, x_N)$ with $x_0, x_1, \ldots, x_N \in \mathbb{R}^n$ being column vectors. This upper index notation is well-defined for all integers $0 \leq k \leq N$. Additionally, we allow the upper index "$[-1]$", which should be read as "none". For example, if we write an expression like

$$\varphi_k(x^{[k-1]}) \;, \; k \in \{0, \ldots, N\} \;,$$

this means that the k-th element of the function sequence $\varphi_0, \varphi_1, \ldots, \varphi_N$ depends on the first k columns of x if $k \geq 1$, but the function φ_0 depends on *none* of the elements of x. Similarly, the lower index notation

$$x_{[k]} = (x_k, x_{k+1}, \ldots, x_N)$$

is used to denote the sub-matrix that consists of all columns of x with index larger than or equal to k. Notice that throughout this paper x denotes a state sequence of a discrete-time system, whose index typically runs from 0 to N. The associated control and disturbance sequences are denoted by u and w and have one element less. For example,

$$u_{[k]} = (u_k, u_{k+1}, \ldots, u_{N-1}) \quad \text{and} \quad w_{[k]} = (w_k, w_{k+1}, \ldots, w_{N-1})$$

denotes the sequences of controls and disturbances from k to $N-1$. Analogous to the upper index notation, we allow an "index overflow", i.e., the lower index "$_{[N]}$" should be read as "none".

2. Preliminaries. This paper concerns the design of control laws for nonlinear discrete time systems of the form

(2.1)
$$\begin{aligned} x_{k+1} &= f(x_k, u_k) + w_k \\ \eta_k &= h(x_k) + v_k \ . \end{aligned}$$

Here, $x_k \in \mathbb{R}^{n_x}$ denotes the state, $u_k \in \mathbb{R}^{n_u}$ the control input, and $w_k \in \mathbb{R}^{n_x}$ the process noise at time $k \in \mathbb{Z}$. The variable $\eta_k \in \mathbb{R}^{n_y}$ denotes the output measurement and $v_k \in \mathbb{R}^{n_h}$ the measurement error. The functions

$$f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_x \times n_x} \quad \text{and} \quad h : \mathbb{R}^{n_x} \to \mathbb{R}^{n_h}$$

denote the right-hand side function of the discrete-time dynamic system as well as the measurement function. Moreover, the stage- and terminal cost are denoted by

$$l : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \to \mathbb{R} \quad \text{and} \quad m : \mathbb{R}^{n_x} \to \mathbb{R} \ .$$

Throughout this paper, the following blanket assumption is used.

ASSUMPTION 1. *The functions f, h, l, and m are three times continuously differentiable with respect to all arguments.*

At many places in the paper, shorthands are used to denote derivatives of the functions f, h, and l. For the first order derivatives the following shorthands are used

$$\begin{aligned} A_k &= \frac{\partial f(x_k, u_k)}{\partial x}, & B_k &= \frac{\partial f(x_k, u_k)}{\partial u}, & C_k &= \frac{\partial h(x_k)}{\partial x}, \\ q_k &= \nabla_x l(x_k, u_k) \ . \end{aligned}$$

Notice that this notation will only be used, if it is clear from the context at which point the derivatives of the functions f, h, and l are evaluated. Similarly, the associated second order derivatives of these functions are denoted by

$$\begin{aligned} K_k &= \nabla_{xx}^2 f(x_k, u_k), & L_k &= \nabla_{ux}^2 f(x_k, u_k), & M_k &= \nabla_{uu}^2 f(x_k, u_k), \\ Q_k &= \nabla_{xx}^2 l(x_k, u_k), & R_k &= \nabla_{uu}^2 l(x_k, u_k), & S_k &= \nabla_{xu}^2 l(x_k, u_k) \ . \end{aligned}$$

All technical derivations are based on the following assumptions about the process noise and measurement errors.

ASSUMPTION 2. *The process noise variables w_k and the measurement errors v_k, $k \in \mathbb{N}$, are independent random variables, whose expectation values and variances $W \succeq 0$ and $V \succeq 0$ are given,*

$$\begin{aligned}
\mathbb{E}\{w_k\} &= 0, & \mathbb{E}\{w_k w_k^{\mathsf{T}}\} &= W \\
\mathbb{E}\{v_k\} &= 0, & \mathbb{E}\{v_k v_k^{\mathsf{T}}\} &= V.
\end{aligned}$$

ASSUMPTION 3. *The process noise variables w_k and the measurement errors v_k have bounded support such that $\|v_k\|_2 \leq \gamma$ as well as $\|w_k\|_2 \leq \gamma$ for a given radius $\gamma > 0$ such that*

$$\|V\|_2 \leq \gamma^2 \quad \text{as well as} \quad \|W\|_2 \leq \gamma^2.$$

Notice that Assumptions 2 and 3 together allow us to compute second order moment expansions of nonlinear functions. For example, if $\varphi : \mathbb{R}^{n_x} \to \mathbb{R}$ is a three times continuously differentiable function of w_k with $\varphi(0) = 0$, we have [45]

$$\mathbb{E}\{\varphi(w_k)\} = \mathbb{E}\left\{ \varphi(0) + \nabla\varphi(0)w_k + \frac{1}{2}w_k^{\mathsf{T}}(\nabla^2\varphi(0))w_k + \mathbf{O}\left(\gamma^3\right) \right\}$$

$$= \mathbb{E}\left\{ \frac{1}{2}\mathrm{Tr}\left(\nabla^2\varphi(0)w_k w_k^{\mathsf{T}}\right) \right\} + \mathbf{O}\left(\gamma^3\right).$$

This equation can also be written in the form

$$\text{(2.2)} \qquad \mathbb{E}\{\varphi(w_k)\} = \frac{1}{2}\mathrm{Tr}(\nabla^2\varphi(w')W) + \mathbf{O}\left(\gamma^3\right),$$

which holds independently of the choice of the point w', at which the derivative $\nabla^2\varphi$ is evaluated, as long as $\|w'\|_2 = \mathbf{O}\left(\gamma\right)$. Similarly, the equation

$$\mathbb{E}\{\varphi(w_k)^2\} = \mathbb{E}\left\{ \left(\varphi(0) + \nabla\varphi(0)w_k + \frac{1}{2}w_k^{\mathsf{T}}(\nabla^2\varphi(0))w_k + \mathbf{O}\left(\gamma^3\right) \right)^2 \right\}$$

$$= \nabla\varphi(0)^{\mathsf{T}}W\nabla\varphi(0) + \mathbf{O}\left(\gamma^3\right)$$

with $W = \mathbb{E}\{w_k w_k^{\mathsf{T}}\}$ implies that we have

$$\text{(2.3)} \qquad \mathbb{E}\{\varphi(w_k)^2\} = \nabla\varphi(w')^{\mathsf{T}}W\nabla\varphi(w') + \mathbf{O}\left(\gamma^3\right).$$

Also this result holds independently of the choice of the point w' as long as we ensure that $\|w'\|_2 = \mathbf{O}\left(\gamma\right)$. Similar expansion results hold for three times continuously differentiable vector-valued functions. For example, if $\tilde{\varphi} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_\varphi}$ is a three times continuously differentiable function, we have

$$\mathrm{Var}(\tilde{\varphi}(w_k)) = \mathbb{E}\left\{ (\tilde{\varphi}(w_k) - \mathbb{E}\{\tilde{\varphi}(w_k)\})(\tilde{\varphi}(w_k) - \mathbb{E}\{\tilde{\varphi}(w_k)\})^{\mathsf{T}} \right\}$$

$$\text{(2.4)} \qquad = \nabla\tilde{\varphi}(w')^{\mathsf{T}}W\nabla\tilde{\varphi}(w') + \mathbf{O}\left(\gamma^3\right)$$

for any w' with $\|w'\|_2 = \mathbf{O}(\gamma)$. We will use equations of the form (2.2), (2.3), and (2.4) in this or very similar versions at many places in this paper (without always mentioning this explicitly).

REMARK 2.1. *In practice, random variables are often modelled by Gaussian probability distributions. Unfortunately, Assumption 3 is violated in this case, as Gaussian distributions do not have a bounded support. It is possible to rescue the equations* (2.2) *and* (2.3) *for Gaussian distributions with small variance,* $\|W\|_2 \leq \gamma$, *if we impose more restrictive assumptions on the global properties of the function* φ. *For example, if* φ *has uniformly bounded third derivatives on* $\mathbb{R}^{n_x}$, *the equations* (2.2) *and* (2.3) *hold. However, for general nonlinear functions* φ *the equations* (2.2) *and* (2.3) *are wrong. Therefore, this paper suggests to model the random variables with probability distributions that closely approximate Gaussian distributions, but have bounded support. One might argue, that this is a realistic modeling assumptions when dealing with nonlinear systems, as noise terms are—due to physical limitations—in practice almost always bounded anyhow, which justifies Assumptions 3.* ◇

3. Motivating Example. Let us consider a nonlinear discrete time system of the form (2.1) with dimensions $n_x = n_u = 2$ as well as $n_y = 1$, given by

(3.1)
$$
\left\{
\begin{array}{rcl}
f(x_k, u_k) & = & F(u_k)x_k + \delta u_k \\
h(x_k) & = & Cx_k
\end{array}
\right\}
\quad \text{with} \quad F(u_k) = I + \delta \left(\begin{array}{cc} 1 & 0 \\ Cu_k & -1 \end{array} \right)
$$

and $C = (0, 1)$. Here, $\delta \ll 1$ is a small time step. This is a nonlinear system as the system matrix $F(u_k)$ depends on the control input u_k. The stage cost is in this example assumed to be a tracking term of the form

(3.2) $\quad l(x_k, u_k) = \delta \left(x_k^\mathsf{T} Q x_k + u_k^\mathsf{T} R u_k \right) \quad \text{with} \quad R = I \quad \text{and} \quad Q = \left(\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right).$

It can be checked easily that this system is unstable in open-loop mode. Since only the second state component, Cx_k, can be measured, the only way to gather information about the first component is to choose $u \neq 0$ such that the first state component has an influence on the propagation of the second state component via the term Cu_k in the lower left corner of the system matrix $F(u_k)$. If we choose u such that $Cu_k = 0$ for all k, the first component of the state vector x cannot be observed. Unfortunately, a standard tracking MPC controller, whose aim is to minimize the stage cost l under the assumption that full-state measurements are available, will choose $Cu_k = 0$ for all k. Thus, the above nonlinear system cannot be stabilized if the MPC controller does not cooperate with the state estimator. For example, a naive cascade consisting of an extended Kalman filter (or MHE) and MPC leads to unstable closed-loop trajectories. The goal of this paper is to develop a receding horizon controller, which resolves this problem while maintaining economic performance.

4. Model Predictive Control. This section briefly reviews the main idea of model predictive control (MPC) and its associated cost-to-go function [49]. For this aim, we introduce for all $i \in \{0, 1, \ldots, N\}$ the notation

(4.1)
$$
\begin{aligned}
J_i(x_i, w_{[i]}) \quad = \quad & \min_{x,u} \quad \sum_{k=i}^{N-1} l(x_k, u_k) + m(x_N) \\
& \text{s.t.} \quad x_{k+1} = f(x_k, u_k) + w_k , \quad k \in \{i, \ldots, N-1\}
\end{aligned}
$$

to denote the parametric optimal control problem that is associated with the nonlinear system f, stage cost l and terminal cost m. The function $J_i : \mathbb{R}^{n_x} \times \mathbb{R}^{n_x \times (N-i)} \to \mathbb{R}$ is called the cost-to-go function that is associated with the time horizon $\{i, i+1, \ldots, N\}$. Here, $w_{[i]} = (w_i, \ldots, w_{N-1})$ denotes the last $(N-i)$ elements of the disturbance sequence as explained in Section 1.2. Notice that the last cost-to-go function, $J_N = m$, is equal to the Mayer term m. The standard implementation of an MPC loop can be outlined as follows.

1. Wait for an estimate y_0 of the current state.
2. Solve Problem (4.1) for $i = 0$, $x_0 = y_0$, and $w = (w_0, w_1, \ldots, w_{N-1}) = 0$.
3. Send the first element u_0 of the optimal control sequence to the real process.
4. Shift the time index and continue with Step 1.

Notice that MPC is a certainty equivalent controller. This means that the future control input is optimized as if there was no uncertainty, $w = 0$, and as if the state estimate $x_0 = y_0$ was accurate. Comments on how to choose the terminal cost m and how to ensure stability of MPC can be found in [14, 40, 49].

REMARK 4.1. *In the following discussion control and/or state constraints are not analyzed explicitly, as this article focuses on how to design controllers for the case that the function f, l, and m are nonlinear. The considerations in the following sections can, however, be generalized for the case that control constraints are present, as explained at the end of Section 7. State constraints are beyond the scope of this paper.* ◇

In the following, we outline the main idea of *dynamic programming* [4, 10]. For this aim, we introduce the auxiliary functions

$$(4.2) \qquad G_k(x_k, u_k, w_{[k]}) = J_{k+1}(f(x_k, u_k) + w_k, w_{[k+1]}) + l(x_k, u_k) \,,$$

which are defined for all $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ and all $w_{[k]} \in \mathbb{R}^{n_x \times (N-k)}$, and all $k \in \{0, 1, \ldots, N-1\}$. The cost-to-go function sequence $J_0, J_1, \ldots, J_N$ satisfies a functional recursion, which is known under the name dynamic programming. It starts with $J_N(\cdot) = m(\cdot)$ and iterates backwards

$$(4.3) \qquad\qquad J_k(\cdot, w_{[k]}) = \min_{u_k} \ G_k(\cdot, u_k, w_{[k]})$$

for all $k \in \{0, 1, \ldots, N-1\}$. Notice that this functional recursion is very useful for analyzing the properties of MPC [49], although numerical implementations of MPC typically solve problem (4.1) directly instead of using dynamic programming. Assumption 1 implies that the function $J_N = m$ as well as the function G_{N-1} are three times continuously differentiable. The following proposition introduces a regularity condition under which this property is inherited by all cost-to-go functions through the dynamic programming recursion.

PROPOSITION 4.1. *Let Assumption 1 be satisfied and let $k \in \{0, \ldots, N-1\}$ be given. If the function J_{k+1} is locally three times continuously differentiable in a neighborhood of $(x_k, 0)$, then the function G_k is locally three times continuously differentiable in all arguments. Moreover, if the minimizer $\mu_k(x_k)$ of problem (4.3)*

is unique and satisfies the first order necessary and second order sufficient optimality conditions,

$$(4.4) \qquad\qquad \nabla_u G_k(x_k, \mu_k(x_k), 0) = 0$$

$$(4.5) \qquad\qquad and \quad G_k^{uu} = \nabla_{uu}^2 G_k(x_k, \mu_k(x_k), 0) \succ 0$$

for all x_k, then the functions J_k is locally three times continuously differentiable in a neighborhood of $(x_k, 0)$.

Proof. The first statement follows trivially from the definition of G_k, as the composition and sum of three times continuously differentiable functions remains three times continuously differentiable. The second statement is a well-known result from the field of parametric nonlinear programming [44], which can be established by applying the implicit function theorem to the first order optimality condition (4.4).◇

Notice that Proposition 4.1 can be applied recursively, in order to show that the cost-to-go functions J_k as well as the functions G_k are—under the regularity assumptions from Proposition 4.1—for all k locally three times differentiable. In this context, the word "locally" means that the functions J_k and G_k are three times continuously differentiable for small disturbance sequences $w_{[k]}$. This is sufficient, as we are in this paper interested in analyzing the influence of small disturbances, i.e., for the case that Assumption 3 is satisfied for a sufficiently small $\gamma > 0$.

ASSUMPTION 4. *We assume that the regularity conditions from Proposition 4.1 are satisfied recursively for all $k \in \{N-1, N-2, \ldots, 0\}$ such that all cost-to-go functions J_k are locally three-times continuously differentiable for all $k \in \{0, \ldots, N\}$.*

In the following, we introduce a shorthand notation for the first and second order derivatives of the cost-to-go functions J_k,

$$p_k = \nabla_x J_k(x_k, 0) \quad and \quad P_k = \nabla_x^2 J_k(x_k, 0) .$$

Notice that the sequence $\Omega_k = (p_k, P_k)$ satisfies an algebraic Riccati recursion [9] of the form

$$(4.6) \quad \forall k \in \{0, \ldots, N-1\}, \quad \Omega_k = \mathcal{F}_\Omega(x_k, u_k, \Omega_{k+1}) \quad and \quad \Omega_N = \mathcal{M}(x_N) .$$

Here, the function

$$\mathcal{M}(x_N) = \left(\nabla_x J_N(x_N), \nabla_x^2 J_N(x_N) \right) = \left(\nabla_x m(x_N), \nabla_x^2 m(x_N) \right) \in \mathbb{R}^{(n_x+1) \times n_x}$$

is for all $x_N \in \mathbb{R}^n$ given by the associated first and second order derivative of the Mayer term m. The expression for the right-hand side function $\mathcal{F}_\Omega$ can be worked out explicitly by differentiating the recursion (4.3) twice. This yields

$$(4.7) \quad \mathcal{F}_\Omega(x_k, u_k, \Omega_{k+1}) = \left(A_k^\mathsf{T} p_{k+1} + q_k \ , \ G_k^{xx} - G_k^{xu} (G_k^{uu})^{-1} G_k^{ux} \right) \in \mathbb{R}^{(n_x+1) \times n_x}.$$

Here, $G_k^{xx}, G_k^{ux} = (G_k^{xu})^\mathsf{T}$, and G_k^{uu} are shorthands for the second order derivatives of the function G_k with respect to x and u,

$$(4.8) \qquad \begin{array}{rcl} G_k^{xx} & = & Q_k + A_k^\mathsf{T} P_{k+1} A_k + K_k \cdot p_{k+1} \\ G_k^{ux} & = & S_k + B_k^\mathsf{T} P_{k+1} A_k + L_k \cdot p_{k+1} \\ G_k^{uu} & = & R_k + B_k^\mathsf{T} P_{k+1} B_k + M_k \cdot p_{k+1} . \end{array}$$

Recall that the shorthands A_k, B_k, K_k, L_k, M_k, Q_k, S_k, and R_k denote first and second order derivatives of the functions f and l with respect to states and controls as defined in Section 2. The first two arguments of the function $\mathcal{F}_\Omega(x_k, u_k, \Omega_{k+1})$ on the left hand side of equation (4.7) indicate that all derivatives are evaluated at the point (x_k, u_k), although this dependence is not highlighted explicitly in the right-hand side of equation (4.7). Finally, a second order expansion of the cost-to-go functions is given by

(4.9) $\quad J_k(x, 0) = J_k(x_k, 0) + p_k^\mathsf{T}(x - x_k) + (x - x_k)^\mathsf{T} P_k(x - x_k) + \mathbf{O}\left(\|x - x_k\|^3\right)$

for all $k \in \{0, 1, \ldots, N\}$.

REMARK 4.2. *If f is jointly affine in (x, u) while l and m are quadratic forms, we have $K_k = 0$, $L_k = 0$, and $M_k = 0$, the functions J_k are quadratic forms in x, and the second order expansion (4.9) is exact. In this case, the algebraic Riccati recursion (4.6) for the sequence $\Omega_k = (p_k, P_k)$ has the form*

$$p_k = A_k^\mathsf{T} p_{k+1} + q_k$$
$$P_k = Q_k + A_k^\mathsf{T} P_{k+1} A_k - \left(S_k + B_k^\mathsf{T} P_{k+1} A_k\right)^\mathsf{T} \left(R_k + B_k^\mathsf{T} P_{k+1} B_k\right)^{-1} \left(S_k + B_k^\mathsf{T} P_{k+1} A_k\right)$$

and the corresponding MPC controller is equivalent to LQR control [9].   ◇

5. State Estimation. This section briefly reviews the extended Kalman filter (EKF) [52] and its relation to moving horizon estimation (MHE) [49]. Recall that $\eta_k = h(x_k) + v_k$ denotes the measurements at time k. The EKF proceeds by maintaining a state estimate $y_k \in \mathbb{R}^{n_x}$ and a variance $\Sigma_k \in \mathbb{R}^{n_x \times n_x}$ by performing updates of the form

(5.1) $$y_{k+1} = \mathcal{F}_y(x_k, u_k, y_k, \eta_k) \,,$$

(5.2) $$\Sigma_{k+1} = \mathcal{F}_\Sigma(x_k, u_k, \Sigma_k)$$

where (x_k, u_k) denotes the point at which the nonlinear functions f and h are linearized. In detail, the right-hand side functions $\mathcal{F}_y$ and $\mathcal{F}_\Sigma$ of the EKF are given by the explicit expressions [52]

$$\mathcal{F}_y(x_k, u_k, y_k, \eta_k) = f\left(y_k - \Sigma_k C_k^\mathsf{T} \left(C_k \Sigma_k C_k^\mathsf{T} + V\right)^\dagger (\eta_k - h(y_k)), u_k\right) \,,$$

$$\mathcal{F}_\Sigma(x_k, u_k, \Sigma_k) = A_k \left(\Sigma_k - \Sigma_k C_k^\mathsf{T} \left(C_k \Sigma_k C_k^\mathsf{T} + V\right)^\dagger C_k \Sigma_k\right) A_k^\mathsf{T} + W \,.$$

The most common variant of the EKF evaluates the derivatives

$$A_k = \frac{\partial f(x_k, u_k)}{\partial x} \quad \text{and} \quad C_k = \frac{\partial h(x_k)}{\partial x}$$

at the current state estimate $x_k = y_k$. However, in the following, we keep our notation general remarking that the first two arguments of the functions $\mathcal{F}_y$ and $\mathcal{F}_\Sigma$ indicate at which point the first order derivatives A_k and C_k of the functions f and h are computed. This notation is needed, if we want to use the extended Kalman in order to estimate how accurate our future state estimates will be without knowing the measurements yet. In such a situation, we can compute the iterates Σ_k of the EKF by evaluating the recursion (5.2) at predicted states x_k in place of the state estimates y_k,

which are in this case not available yet. This is possible, as the function $\mathcal{F}_\Sigma$ does not depend on η_k. In the following, we denote the true state sequence by $z_0, z_1, z_2, \ldots \in \mathbb{R}^{n_x}$. This sequence is defined by the recursion

$$(5.3) \qquad \forall k \in \mathbb{N}, \qquad z_{k+1} = f(z_k, u_k, w_k) \quad \text{with} \quad z_0 = x_0^* .$$

Notice that this recursion depends on the disturbance sequence $w_0, w_1, w_2, \ldots$ as well as the true state x_0^* of the system at time 0, which are all unknown.

LEMMA 5.1. *Let Assumptions 1, 2, and 3 be satisfied, and let the sequences $y_0, y_1, y_2, \ldots$ and $\Sigma_0, \Sigma_1, \Sigma_2, \ldots$ be given by the extended Kalman filter recursions (5.1) and (5.2), where the linearization points x_k satisfy $x_k = z_k + \mathbf{O}\left(\gamma\right)$ while the sequence u_k is given and exactly known. If the initial state estimate y_0 and the initial variance Σ_0 satisfy $\Sigma_0 = \mathbf{O}\left(\gamma^2\right)$ as well as*

$$\mathbb{E}\{y_0\} = x_0^* + \mathbf{O}\left(\gamma^2\right) \quad and \quad \mathbb{E}\{(y_0 - x_0^*)(y_0 - x_0^*)^\mathsf{T}\} = \Sigma_0 + \mathbf{O}\left(\gamma^3\right) ,$$

then the following statements hold on any finite time horizon, $k \in \{0, 1, \ldots, N\}$.

1. *We have $\Sigma_k = \mathbf{O}\left(\gamma^2\right)$.*
2. *The expectation of the state estimate satisfies $\mathbb{E}\left\{y_k\right\} = z_k + \mathbf{O}\left(\gamma^2\right)$.*
3. *The variance of the state estimate satisfies*

$$\mathrm{Var}(y_k - z_k) = \mathbb{E}\left\{(y_k - z_k)(y_k - z_k)^\mathsf{T}\right\} + \mathbf{O}\left(\gamma^3\right) = \Sigma_k + \mathbf{O}\left(\gamma^3\right) .$$

Proof. The statement of Lemma 5.1 is—at least in very similar versions—well known in the literature [52]. The main idea is to use induction. Firstly, all three statements of the Lemma are true for $k = 0$. Next, under the assumption that these three statements are satisfied for a given k, they can be established for $k + 1$: if $\Sigma_k = \mathbf{O}\left(\gamma^2\right)$ for a given k, then we also have

$$\Sigma_{k+1} = F_\Sigma(x_k, u_k, \Sigma_k) = \mathbf{O}\left(\|\Sigma_k\|_2\right) + \mathbf{O}\left(\|W\|_2\right) = \mathbf{O}\left(\gamma^2\right) ,$$

i.e., the first statement holds for all $k \in \{-N, \ldots, 0\}$. The second and third statement follows from the fact the EKF computes the first and second order moments exactly, if the dynamic system f and the measurement function h are affine in x. Thus, we can apply the Taylor expansion techniques from Section 2 to bound the contribution of higher order terms in order to show that the second and third statement also hold for $k + 1$. The details of this argument can be found in [23]. ◇

REMARK 5.1. *The statement of Lemma 5.1 relies on the assumption that the horizon N is finite, but the statement of this lemma is wrong in general for infinite horizons without further observability assumptions, as the EKF updates are divergent in general.* ◇

REMARK 5.2. *Lemma 5.1 can also be applied for the case that the state estimates and variances are computed with a moving horizon estimator (MHE) [49], whose arrival cost is computed by extended Kalman filter updates. This leads to a refinement of the state estimates for nonlinear system. However, MHE with suitable arrival cost updates is in a first order approximation equivalent to EKF. A proof of this statement can be found in Chapter 3 of [23].* ◇

6. Expected Loss of Optimality. In Section 3 we have discussed an example, where certainty equivalent MPC leads to a sub-optimal performance, as it tries to bring the system to a non-observable steady-state. The goal of this section is to analyze in detail how much performance is actually lost when running a certainty equivalent MPC to control a general nonlinear system in the presence of disturbances and inexact state estimates. Let x_0^* denote the true (but unknown) value of the state at time $t = 0$. If we would know the future disturbance sequence $w^* = \left(w_0^*, w_1^*, \ldots, w_{N-1}^*\right)$ in advance, we could compute the optimal input sequence $u^* = \left(u_0^*, u_1^*, \ldots, u_{N-1}^*\right)$ as the solution of the optimal control problem

(6.1)
$$J^* = J_0(x_0^*, w_{[0]}^*) \quad = \quad \min_{x,u} \quad \sum_{k=0}^{N-1} l(x_k, u_k) + m(x_N)$$
$$\text{s.t.} \quad \begin{cases} \forall k \in \{0, \ldots, N-1\}, \\ x_{k+1} = f(x_k, u_k) + w_k^* \\ x_0 = x_0^* . \end{cases}$$

Of course, in practice, we neither know the true initial value x_0^* nor can predict the future disturbance sequence w^*. Thus, the optimal input sequence u^* cannot be computed. Recall that the control law, which is associated with the certainty equivalent MPC controller (4.1), is given by

$$\mu_k(\cdot) \quad = \quad \operatorname*{argmin}_{u_k} G_k(\cdot, u_k, 0) ,$$

where the functions G_k have been defined in equation (4.2) in Section 4. The true state of the closed loop MPC system at time k, denoted by $\xi_k^*(y^{[k-1]})$, is a function of the state estimate sequence

$$y^{[k-1]} = (y_0, \ldots, y_{k-1}) .$$

It can be obtained recursively as the solution of the certainty equivalent MPC recursion

(6.2)
$$\begin{aligned} \xi_0^* &= x_0^* \\ \xi_1^*(y^{[0]}) &= f(\xi_0^*, \mu_0(y_0)) + w_0^* \\ \xi_2^*(y^{[1]}) &= f(\xi_1^*(y^{[0]}), \mu_1(y_1)) + w_1^* \\ &\vdots \\ \xi_N^*(y^{[N-1]}) &= f(\xi_{N-1}^*(y^{[N-2]}), \mu_{N-1}(y_{N-1})) + w_{N-1}^* . \end{aligned}$$

In the following, we denote with

(6.3)
$$J_{\mathrm{cl}}^*(y^{[N-1]}) \quad = \quad \sum_{k=0}^{N-1} l(\xi_k^*(y^{[k-1]}), \mu_k(y_k)) + m(\xi_N^*(y^{[N-1]}))$$

the actual performance of the MPC controller. If uncertainties are present, the performance of the MPC controller can be expected to be worse than the best possible performance, given by

(6.4)
$$J^* \quad = \quad \sum_{i=0}^{N-1} l(x_i^*, u_i^*) + m(x_N^*) = J_{\mathrm{cl}}^*((x_0^*, x_1^*, \ldots, x_{N-1}^*)) .$$

We call the associated difference, given by

$$\Delta^*(y) = J_{\mathrm{cl}}^*(y) - J^* ,$$

the "loss of optimality". In the following, we use the notation

$$(6.5) \quad \Delta_k^*(y^{[k]}) = G_k(\xi_k^*(y^{[k-1]}), \mu_k(y_k), w_{[k]}^*) - G_k(\xi_k^*(y^{[k-1]}), \mu_k(\xi_k^*(y^{[k-1]})), w_{[k]}^*)$$

to denote the loss of optimality that is associated with the k-th suboptimal decision. The lemma below establishes an important relation between the function Δ^* and the function sequence $\Delta_0^*, \ldots, \Delta_{N-1}^*$.

LEMMA 6.1. *Let Assumption 1 be satisfied. We have* $\Delta^*(y) = \sum_{k=0}^{N-1} \Delta_k^*(y^{[k]})$.

Proof. Recall that the functions G_k are defined by equation (4.2). Together with equation (6.2), this definition implies

$$G_k(\xi_k^*(y^{[k-1]}), \mu_k(y_k), w_{[k]}^*) = J_{k+1}(f(\xi_k^*(y^{[k-1]}), \mu_k(y_k)) + w_{[k]}^*, w_{[k+1]}^*)$$

$$+ l(\xi_k^*(y^{[k-1]}), \mu_k(y_k))$$

$$(6.6) \qquad\qquad = J_{k+1}(\xi_{k+1}^*(y^{[k]}), w_{[k+1]}^*) + l(\xi_k^*(y^{[k-1]}), \mu_k(y_k)) \ .$$

Moreover, since $\mu_k(\xi_k^*(y^{[k-1]}))$ is, by definition, a minimizer of the function $G_k(\xi_k^*(y^{[k-1]}), \cdot, w_{[k]}^*)$ the dynamic programming equation (4.3) yields

$$(6.7) \qquad G_k(\xi_k^*(y^{[k-1]}), \mu_k(\xi_k^*(y^{[k-1]})), w_{[k]}^*) = J_k(\xi_k^*(y^{[k-1]}), w_{[k]}^*) \ .$$

Subtracting equations (6.6) and (6.7) yields

$$\Delta_k^*(y^k) = J_{k+1}(\xi_{k+1}^*(y^{[k]}), w_{[k+1]}^*) - J_k(\xi_k^*(y^{[k-1]}), w_{[k]}^*) + l(\xi_k^*(y^{[k-1]}), \mu_k(y_k))$$

for all $k \in \{0, 1, \ldots, N-1\}$. Consequently, the sum over the terms $\Delta_k^*(y^k)$ turns out to be a telescoping series, which can be simplified as follows:

$$\sum_{k=0}^{N-1} \Delta_k^*(y^{[k]}) = J_N(\xi_N^*(y^{[N-1]})) - J_0(\xi_0^*, w_{[0]}^*) + \sum_{k=0}^{N-1} l(\xi_k^*(y^{[k-1]}), \mu_k(y_k))$$

$$= \sum_{k=0}^{N-1} l(\xi_k^*(y^{[k-1]}), \mu_k(y_k)) + m(\xi_N^*(y^{[N-1]})) - J_0(x_0^*, w_{[0]}^*)$$

$$(6.8) \qquad\qquad \overset{(6.3)}{=} J_{\text{cl}}^*(y) - J^* = \Delta^*(y) \ .$$

This is the statement of the lemma. $\diamond$

The performance of an MPC controller that operates under inexact state measurements and process noise can never have a better performance than the utopian controller from equation (6.1), which knows the true current state and all future process disturbances. Thus, the loss of optimality function Δ^* must be non-negative. The following corollary summarizes this result and provides a formal proof.

COROLLARY 6.2. *Let Assumption 1 be satisfied. We have* $\Delta^*(y) \geq 0$ *for all measurement sequences* $y = y^{[N-1]}$.

Proof. Definition (6.5) implies that the functions Δ_k are non-negative, as $\mu_k(\xi_k^*(y^{[k-1]}))$ is a minimizer of the function $G_k(\xi_k^*(y^{[k-1]}), \cdot, w_{[k]}^*)$. Thus Lemma 6.1 implies that the function Δ^* is non-negative, too.   ◇

In the following, we focus on the derivation of a tractable second order approximation of the expected loss of optimality, given by

$$\mathbb{E}\{\Delta^*(y)\} = \sum_{k=0}^{N-1} \mathbb{E}\left\{\Delta_k^*(y^{[k]})\right\} \ .$$

Here, the expected value is computed under the assumption that the state estimates y are computed by the EKF (or locally equivalent MHE) that is described in Section 5. On the first view, we might expect that the computation of such a second order expansion is rather cumbersome and computationally expensive, as the state estimates y_k are in general correlated in time. Nevertheless, the following theorem provides a surprisingly simple recursion technique for computing a second order approximation of the expected loss of optimality requiring $\mathbf{O}(N)$ operations only. It is based on the following notation.

- The matrices $\Omega_k = [p_k, P_k]$ denote the solution of the backward Riccati recursion

  (6.9) $\Omega_k = \mathcal{F}_\Omega(x_k, u_k, \Omega_{k+1})$ with $\Omega_N = \mathcal{M}(x_N)$

  for $k \in \{0, \dots, N-1\}$, which has been introduced in Section 4.

- We use the shorthand notation

  $$\Phi_k = G_k^{xu}(G_k^{uu})^{-1} G_k^{ux}$$
  $$= \left(S_k + B_k^\mathsf{T} P_{k+1} A_k + L_k \cdot p_{k+1}\right)^\mathsf{T} \cdot \left(R_k + B_k^\mathsf{T} P_{k+1} B_k + M_k \cdot p_{k+1}\right)^{-1}$$
  $$\cdot \left(S_k + B_k^\mathsf{T} P_{k+1} A_k + L_k \cdot p_{k+1}\right) .$$

  Notice that the matrices Φ_k need to be computed anyhow during the evaluation of the backward recursion (6.9), as the recursion for the matrices P_k can be written in the explicit form

  (6.10) $P_k = \underbrace{\left(Q_k + A_k^\mathsf{T} P_{k+1} A_k + K_k \cdot p_{k+1}\right)}_{G_k^{xx}} - \Phi_k \ .$

- Finally, the matrices Σ_k are given by the forward Riccati recursion

  (6.11) $\Sigma_{k+1} = \mathcal{F}_\Sigma(x_k, u_k, \Sigma_k)$ with $\Sigma_0 = \hat{\Sigma}_0$

  where $\hat{\Sigma}_0 = \mathbb{E}\{(y_0 - x_0^*)(y_0 - x_0^*)^\mathsf{T}\} + \mathbf{O}(\gamma^3)$ denotes an approximation of the variance of the initial state estimate y_0 as introduced in Section 5.

Theorem 6.3 in combination with Corollary 6.4 are one of the core technical contributions of this paper.

THEOREM 6.3. *Let Assumptions 1, 2, 3, and 4 be satisfied, then we have*

$$\mathbb{E}\left\{\Delta_k^*(y^{[k]})\right\} = \frac{1}{2}\mathrm{Tr}(\Phi_k \Sigma_k) + \mathbf{O}(\gamma^3)$$

for all $k \in \{0, \ldots, N-1\}$. Here, the matrices Ω_k, Φ_k, and Σ_k are computed as outlined above and under the assumption that the linearization point (x_k, u_k), at which the forward and backward Riccati recursions for the sequences Ω_k and Σ_k as well as the matrices Φ_k are evaluated, satisfies $(x_k, u_k) = (y_k, \mu_k(y_k)) + \mathbf{O}(\gamma)$.

Proof. In order to understand the following proof it is helpful to recall that the state estimates $y_1, \ldots, y_k$ are correlated random variables. Thus, if we compute the expectation with respect to the most recent state estimate y_k first,

$$(6.12) \qquad \mathbb{E}\left\{\Delta_k^*(y^{[k]})\right\} = \underset{y^{[k-1]}}{\mathbb{E}}\left\{\underset{y_k}{\mathbb{E}}\left\{\Delta_k^*(y^{[k]})\right\}\right\},$$

we should keep in mind that $\underset{y_k}{\mathbb{E}}\left\{\Delta_k^*(y^{[k]})\right\}$ denotes the conditional expectation value of $\Delta_k^*(y^{[k]})$ over the variable y_k, which does depend on the values of the previous state estimates $y^{[k-1]} = (y_0, \ldots, y_{k-1})$. Since μ_k is an optimizing control law, the first order necessary optimality condition

$$\left.\frac{\partial G_k(\xi_k^*(y^{[k-1]}), u, w_{[k]}^*)}{\partial u}\right|_{u=\mu_k(\xi_k^*(y^{[k-1]}))} = 0$$

must be satisfied. Thus, it follows from equation (6.5) that the equations

$$(6.13) \qquad 0 = \Delta_k^*(y^{[k]})\Big|_{y_k=\xi_k^*(y^{[k-1]})} \qquad \text{and} \qquad 0 = \left.\frac{\partial}{\partial y_k}\Delta_k^*(y^{[k]})\right|_{y_k=\xi_k^*(y^{[k-1]})}$$

hold independently of the choice of $y^{[k-1]}$. Next, we introduce the notation

$$\Gamma_k(y^{[k-1]}) = \left.\nabla_{y_k}^2 \Delta_k^*(y^{[k]})\right|_{y_k=\xi_k^*(y^{[k-1]})}$$

to denote the Hessian matrix of the function Δ_k^*. This Hessian matrix can be worked out by applying the implicit function theorem in combination with the chain rule,

$$\Gamma_k(y^{[k-1]}) = \left.\frac{\partial^2 G_k(x, u, w^*)}{\partial x \partial u}\right|_{\substack{u=\mu_k(\xi_k^*(y^{[k-1]})) \\ x=\xi_k^*(y^{[k-1]})}} \cdot \left(\left.\frac{\partial^2 G_k(x, u, w^*)}{\partial u \partial u}\right|_{\substack{u=\mu_k(\xi_k^*(y^{[k-1]})) \\ x=\xi_k^*(y^{[k-1]})}}\right)^{-1}$$

$$\cdot \left.\frac{\partial^2 G_k(x, u, w^*)}{\partial u \partial x}\right|_{\substack{u=\mu_k(\xi_k^*(y^{[k-1]})) \\ x=\xi_k^*(y^{[k-1]})}} \cdot$$

Fortunately, this equation can be written in the form

$$\Gamma_k(y_0, \ldots, y_{k-1}) = \Phi_k + \mathbf{O}(\gamma)$$

with $\Phi_k = G_k^{xu}(G_k^{uu})^{-1}G_k^{ux}$, since we assume that the evaluation point (x_k, u_k) satisfies

$$(x_k, u_k) = (y_k, \mu_k(y_k)) + \mathbf{O}(\gamma) .$$

Moreover, it follows from the consideration in Section 5 that we have

$$\mathbb{E}\left\{\left(y_k - \xi_k^*(y^{[k-1]})\right)\left(y_k - \xi_k^*(y^{[k-1]})\right)^\mathsf{T}\right\} = \Sigma_k + \mathbf{O}\left(\gamma^3\right) \ .$$

Consequently, by using the second order moment expansion technique from Section 2, the term $\underset{y_k}{\mathbb{E}}\left\{\Delta_k^*(y^{[k]})\right\}$ can be written in the form

$$\underset{y_k}{\mathbb{E}}\left\{\Delta_k^*(y^{[k]})\right\} = \frac{1}{2}\operatorname{Tr}\left(\Phi_k\Sigma_k\right) + \mathbf{O}\left(\gamma^3\right) \ .$$

This equation is very fortunate, since it implies that the conditional expectation $\underset{y_k}{\mathbb{E}}\left\{\Delta_k^*(y^{[k]})\right\}$ is in a second order approximation independent of $y^{[k-1]}$. Thus, the expectation value over all random variables can be evaluated as

$$\mathbb{E}\left\{\Delta_k^*(y^{[k]})\right\} = \underset{y^{[k-1]}}{\mathbb{E}}\left\{\underset{y_k}{\mathbb{E}}\left\{\Delta_k^*(y^{[k]})\right\}\right\} = \tfrac{1}{2}\operatorname{Tr}\left(\Phi_k\Sigma_k\right) + \mathbf{O}\left(\gamma^3\right) \ .$$

This completes the proof of Theorem 6.3. ◇

COROLLARY 6.4. *If the conditions from Theorem 6.3 are satisfied, then the expected loss of optimality is given by*

$$\mathbb{E}\{\Delta^*(y)\} = \frac{1}{2}\sum_{k=0}^{N-1}\operatorname{Tr}(\Phi_k\Sigma_k) + \mathbf{O}\left(\gamma^3\right) \ .$$

Proof. The statement of this corollary is a trivial consequence of Theorem 6.3 and Lemma 6.1. ◇

7. Self-Reflective Model Predictive Control. This section proposes a receding horizon controller, whose objective is to minimize the sum of the certainty equivalent control performance and a second order approximation of the associated expected loss of optimality due to inaccurate state estimates. For this aim, we recall that the second order approximation Σ_k of the variance of the future state estimates satisfies the forward recursion

$$\Sigma_{k+1} = \mathcal{F}_\Sigma(x_k, u_k, \Sigma_k) \quad \text{with} \quad \Sigma_0 = \hat{\Sigma}_0$$

for all $k \in \{0, \ldots, N-1\}$. The associated initial value is assumed to satisfy

$$\hat{\Sigma}_0 = \mathbb{E}\{(y_0 - x_0^*)(y_0 - x_0^*)^\mathsf{T}\} + \mathbf{O}\left(\gamma^3\right) \ ,$$

where y_0 denotes the current state estimate and x_0^* the true state at time $t = 0$. As in the previous section $\Omega_k = [p_k, P_k]$ denotes the state of the backward Riccati recursion (4.6) with the parametric final state $\Omega_N = \mathcal{M}(x_N)$. Motivated by Theorem 6.3, we introduce the notation

$$\mathcal{L}(x_k, u_k, \Sigma_k, \Omega_{k+1}) = \frac{1}{2}\operatorname{Tr}(\Phi_k\Sigma_k)$$

to denote the second order approximation of the expected loss of optimality that is associated with the k-th future decision. Here, the matrix Φ_k is a function of the

linearization point (x_k, u_k) and the state Ω_{k+1} of the backward Riccati recursion. The focus of this section is on the parametric optimal control problem

$$J_i^{\mathrm{SR}}(x_i, \Sigma_i) = \min_{x,u,\Sigma,\Omega} \sum_{k=i}^{N-1} \{l(x_k, u_k) + \mathcal{L}(x_k, u_k, \Sigma_k, \Omega_{k+1})\} + m(x_N)$$

(7.1)

$$\text{s.t.} \quad \begin{cases} \forall k \in \{i, \dots, N-1\}, \\ x_{k+1} = f(x_k, u_k), \\ \Sigma_{k+1} = \mathcal{F}_\Sigma(x_k, u_k, \Sigma_k), \\ \Omega_k = \mathcal{F}_\Omega(x_k, u_k, \Omega_{k+1}), \\ \Omega_N = \mathcal{M}(x_N). \end{cases}$$

The full self-reflective MPC loop can be outlined as follows.

1. Wait for the measurement η_0.
2. Update the current state estimate and variance by using an EKF

$$\hat{y} \leftarrow \mathcal{F}_y(\hat{y}, u_{-1}, \hat{y}, \eta_0) ,$$

$$\hat{\Sigma} \leftarrow \mathcal{F}_\Sigma(\hat{y}, u_{-1}, \hat{\Sigma}) .$$

3. Solve Problem (7.1) for $i = 0$, $x_0 = \hat{y}$, and $\Sigma_0 = \hat{\Sigma}$.
4. Send the solution for u_0 to the real process.
5. Reset the time index and continue with Step 1.

In the following, we denote with $\mu_i^{\mathrm{SR}}(x_i, \Sigma_i)$ the optimal solution of Problem (7.1) for the optimization variable u_i in dependence on x_i and Σ_i. The true closed-loop trajectory is then given by the coupled recursion

$$(7.2) \quad \zeta_{k+1}^*(y^{[k]}) = f(\zeta_k^*(y^{[k-1]}), \mu_k^{\mathrm{SR}}(y_k, \Sigma_k^*(y^{[k-1]}))) + w_k^* \quad \text{with} \quad \zeta_0^* = x_0^*$$

$$(7.3) \quad \Sigma_{k+1}^*(y^{[k]}) = F_\Sigma(y_k, \mu_k^{\mathrm{SR}}(y_k, \Sigma_k^*(y^{[k-1]})), \Sigma_k^*(y^{[k-1]})) \quad \text{with} \quad \Sigma_0^* = \Sigma_0^* .$$

The loss of optimality of the above outlined controller is given by

$$\Delta_{\mathrm{SR}}^*(y^{[N-1]}) = \sum_{k=0}^N l(\zeta_k^*(y^{[k-1]}), \mu_k^{\mathrm{SR}}(y_k, \Sigma_k^*(y^{[k-1]}))) + m(\zeta_N^*(y^{[N-1]})) - J^* .$$

Finally, the expected loss of optimality is denoted by $\mathbb{E}_y \left\{ \Delta_{\mathrm{SR}}^*(y^{[N-1]}) \right\}$. The following corollary is a simple consequence of Theorem (6.3).

COROLLARY 7.1. *Let Assumptions 1, 2, 3, and 4 be satisfied, then the above outlined controller from Problem* (7.1) *is (approximately) self-reflective in the sense that it optimizes the sum of its nominal performance and a second order approximation of its own expected loss of optimality. In other words, if* (x, u, Σ, Ω) *denotes an optimal solution of Problem* (7.1) *for* $i = 0$ *and*

$$\mathbb{E}\{\hat{y}\} = x_0^* + \mathbf{O}\left(\gamma^2\right) \quad and \quad \hat{\Sigma} = \mathbb{E}\{(y_0 - x_0^*)(y_0 - x_0^*)^\mathsf{T}\} + \mathbf{O}\left(\gamma^3\right) ,$$

then we have

$$\mathbb{E}_y \left\{ \Delta_{\mathrm{SR}}^*(y^{[N-1]}) \right\} = \sum_{k=0}^{N-1} \mathcal{L}(x_k, u_k, \Sigma_k, \Omega_{k+1}) + \mathbf{O}\left(\gamma^3\right) .$$

Proof. Notice that Problem (7.1) and Problem (4.1) differ for small γ only by the small (but not insignificantly small) additional term $\mathcal{L}(x_k, u_k, \Sigma_k, \Omega_{k+1})$, which satisfies

$$\mathcal{L}(x_k, u_k, \Sigma_k, \Omega_{k+1}) = \mathbf{O}\left(\gamma^2\right) .$$

Thus, due to the regularity requirement from Assumption 4, this additional term perturbs the control law by terms of order $\mathbf{O}\left(\gamma^2\right)$ only,

$$\mu_k^{\mathrm{SR}}(x, \Sigma) = \mu_k(x) + \mathbf{O}\left(\gamma^2\right)$$

as long as $\Sigma = \mathbf{O}\left(\gamma^2\right)$. An analogous conclusion holds for the state trajectory, which implies

(7.4) $$\mathbb{E}_y\left\{\Delta_{\mathrm{SR}}^*(y^{[N-1]})\right\} = \mathbb{E}_y\left\{\Delta^*(y^{[N-1]})\right\} + \mathbf{O}\left(\gamma^4\right) .$$

We know from Theorem 6.3 that the expected loss of optimality of the certainty equivalent MPC is given by $\mathbb{E}_y\left\{\Delta^*(y^{[N-1]})\right\} = \mathcal{L}(x_k, u_k, \Sigma_k, \Omega_{k+1}) + \mathbf{O}\left(\gamma^3\right)$ and that this equation holds independently of the linearization point (x_k, u_k) as long as we do not move out of a $\mathbf{O}\left(\gamma\right)$ neighborhood of the true optimal state and input sequence. This proves the statement of the corollary. ◇

REMARK 7.1. *Problem (7.1) is not a standard optimal control problem in the sense that the stage cost contains a term that couples the k-th state x_k, Σ_k and the $(k+1)$-th state Ω_{k+1}. However, Problem (7.1) is still band-structured and there exist structure-exploiting solvers for implicit dynamic systems [7], which can be used to solve problems of the form (7.1) with run-time complexity $\mathbf{O}\left(N\right)$.* ◇

REMARK 7.2. *As already mentioned in the introduction, existing methods on persistently exciting MPC controllers such as [31] are similar to the proposed controller in the sense that they also add a term that minimizes the trace or other scalar measures of the predicted variance state Σ_k of future state estimates. However, the main contribution of this paper is that we compute weighting matrices Φ_k such that the penalty terms can be interpreted as the expected loss of optimality of the controller. Notice that not only the variances Σ_k of future state estimates but also the weighting matrices Φ_k may depend on the control input u. If we would only add a term of the form $\mathrm{Tr}(\Sigma_k)$ to the objective, this might lead to more accurate state estimates but it does not necessarily improve the expected control performance, if the associated control decisions lead to an increase of the eigenvalues of the matrix Φ_k.* ◇

REMARK 7.3. *The above analysis can be generalized for the case that control constraints are present. One obvious solution for including such constraints is to relax them and add barrier functions [44] to the Lagrange term l, which penalize constraint violations. If such "soft-constraints" are acceptable, the barrier functions can be chosen to be smooth such that the above analysis can be applied. Otherwise, if Problem (4.1) is augmented with hard constraints on the control input, the cost-to-go functions J_i and, consequently, also the associated functions G_i are only piecewise differentiable. Thus, second order expansions of the expected loss of optimality are*

only possible in the neighborhood of regular points, where no active set changes are expected. If we are at a non-differentiable point of the cost-to-go function, where active set changes have to be taken into account, estimates of the expected loss of optimality can be evaluated by using Monte-Carlo simulations, which are, however, much more expensive than Taylor expansion based second order estimates. Other strategies for dual control in the presence of control constraints may be possible and an interesting research direction, but are beyond the scope of the current paper.

A rigorous analysis of the limit behavior and the associated question about the stability of the proposed self-reflective controller (7.1) is an important open problem, which is, however, beyond the scope of this paper. Notice that a rigorous stability analysis of existing persistently exciting MPC, let alone dual controllers, is at the current status of research equally unavailable. However, there exists a number of articles that address the stability of general economic MPC controllers with and without terminal costs [2, 3, 13, 22, 30, 59]. Also notice that the presented self-reflective MPC scheme does not address the question how the uncertainty affects the states, which is in contrast to tube based robust MPC [47]. An analysis of the robustness properties of self-reflective MPC might be an interesting topic for future research.

8. Case Study. This section presents a numerical case study for a predator-prey system of the form

$$(8.1) \qquad \dot{z}_1(t) = p_1 z_1(t) - p_2 z_1(t) z_2(t) + w_1(t) \, ,$$

$$(8.2) \qquad \dot{z}_2(t) = -p_3 z_2(t) + p_4 z_1(t) z_2(t) + u(t) z_3(t) z_2(t) + w_2(t) \, ,$$

$$(8.3) \qquad \dot{z}_3(t) = -z_3(t) + 0.5 + w_3(t) \, .$$

This model is a slightly modified version of the predator-prey optimal experiment design benchmark problem from [50]. Here, the states z_1 and z_2 can be interpreted as the population densities of a prey and a predator species, respectively. We can use the control input $u(t)$ to either hunt or feed the predator with random success rate $z_3(t) \in [0, 1]$. Notice that the approximately Brownian noise $w_3(t) \in [-0.5, 0.5]$ is assumed to have bounded support such that $z_3(t)$ is for all $t > 0$ in $[0, 1]$ as long as $z_3(0) \in [0, 1]$. The functions $w_1(t)$ and $w_2(t)$ model additional external approximately Brownian noise affecting the population density of the predator and prey, respectively. For simplicity of presentation, we assume in this case study that the parameters $p_1 = p_2 = p_3 = p_4 = 1$ are exactly known. Moreover, we assume that we can only measure the population of the prey,

$$h(z(t)) = z_1(t) \, ,$$

but we have neither direct measurements of z_2 nor of z_3. Now, our nominal objective has the form

$$\int_0^T \left((z_1(t) - 1)^2 + (z_2(t) - 1)^2 + u(t)^2 \right) \, \mathrm{d}t + \underbrace{(z(T) - 1)^\mathsf{T} P_T (z(T) - 1)}_{=m(z(T))} \, ,$$

i.e., we would like to keep the population $(z_1(t), z_2(t))$ close to its steady state $(1, 1)$. Notice that the system is not observable at the steady state, as the corresponding steady-state control input is given by $u_\mathrm{s} = 0$. The weighting matrix

$$P_T = \begin{pmatrix} 2 + \sqrt{2} & 1 + \sqrt{2} \\ 1 + \sqrt{2} & 1 + \sqrt{2} \end{pmatrix}$$
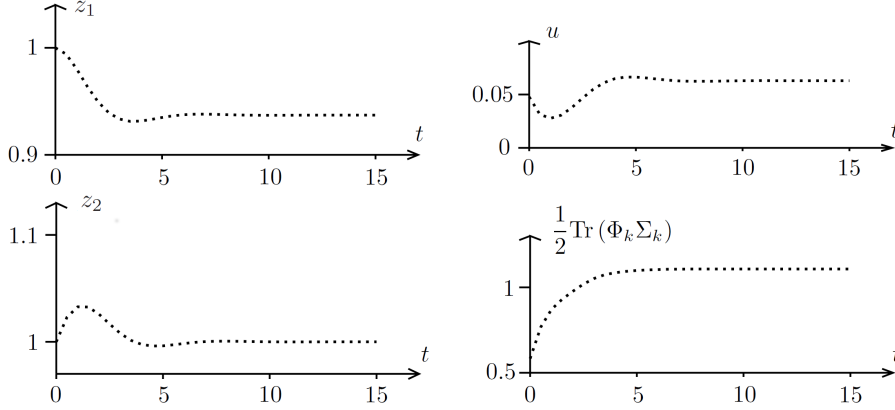
Fig. 1. *LEFT: the closed-loop state trajectories. RIGHT: the closed-loop control input as well as stage-wise expected loss of optimality of the self-reflective controller.*

is constant and chosen in such a way that the terminal cost m can be interpreted as an approximation of the infinite horizon cost. In this example, we choose the prediction horizon $T = 10$. Moreover, we use the software ACADO-Toolkit [28] in order to discretize the above continuous-time system and stage cost with high accuracy such that the associated discretization error can be neglected. The sampling time of the controller is set to $\delta = 0.5$ such that the discrete-time horizon of the associated self-reflective model predictive control problem (7.1) is given by $N = T/\delta = 20$. The matrices V and W are in our case study given by

$$V = \frac{4}{\delta} \quad \text{and} \quad W = \frac{1}{\delta} \operatorname{diag}(\, 0.1, \, 0.1, \, 0.5 \,) \, .$$

Moreover, the closed-loop system is started with

$$\hat{y}_0 = \begin{pmatrix} 1 \\ 1 \\ 0.5 \end{pmatrix} \qquad \text{and} \qquad \hat{\Sigma}_0 = \operatorname{diag}(\, 0.1, \, 0.1, \, 0.5 \,) \, .$$

We simulate the self-reflective MPC closed-loop system without applying process noise or measurement errors (but the controller does not know this). Figure 1 shows the numerical solution for the closed-loop trajectories for the predator and prey population densities as well as the control input and expected loss of optimality of all control decisions in dependence on time (iteration index). Notice that the self-reflective MPC controller does not attempt to track the desired steady state at $(1,1)$. Rather the closed-loop state trajectory for the prey population converges approximately to 0.94. This is in contrast to nominal MPC, which would choose $u = 0$ and track the reference exactly. This trend can also be seen from the closed-loop control profile in the upper right corner of Figure 1, where it becomes apparent that it is optimal to choose a slightly positive control input in order to enforce observability.

9. Conclusions. In this paper a self-reflective model predictive controller has been proposed. Important contributions are Lemma 6.1 and Theorem 6.3 establishing a recursion scheme for computing tractable second order expansions of the expected loss of optimality that is associated with certainty equivalent MPC schemes. This

recursion scheme has been used to develop a model based controller, which is self-reflective in the sense that it minimizes the sum of its nominal performance and a second order approximation of its own expected loss of optimality due to inexact future measurements. Moreover, a small but non-trivial case study has illustrated that the controller is capable of exciting the system in order to improve future state estimate while maintaining economic control performance.

REFERENCES

[1] F. Allgöwer and A. Zheng. *Nonlinear Predictive Control*. Birkhäuser, Progress in Systems Theory, (26), 2000.

[2] R. Amrit, J.B. Rawlings, D. Angeli. Economic optimization using model predictive control with a terminal cost? Annual Reviews in Control, 35(2):178–186, 2011.

[3] D. Angeli, R. Amrit, J.B. Rawlings. On Average Performance and Stability of Economic Model Predictive Control. IEEE Transactions of Automatic Control, 57:1615–1626, 2012.

[4] R.E. Bellman. Dynamic Programming. Princeton University Press, Princeton, 1957.

[5] A. Bemporad and A. Garulli. *Output-feedback predictive control of constrained linear systems via set-membership state estimation.* International Journal of Control, 73(8):655–665, 2000.

[6] A. Bemporad, F. Borrelli, and M. Morari. *Model Predictive Control Based on Linear Programming - The Explicit Solution*, IEEE Transactions on Automatic Control, 47:1974–1985, 2002.

[7] L.T. Biegler and J.B. Rawlings. *Optimization approaches to nonlinear model predictive control*, Proc. 4th International Conference on Chemical Process Control-CPC IV, pp:543–571, 1991.

[8] R.R. Bitmead, M. Gevers, V. Wertz. Adaptive optimal control: The thinking man's gpc. Prentice-Hall, Englewood Cliffs, NJ, ISBN 0-13-013277-2, 1990.

[9] S. Bittanti, P. Colaneri, G. De Nicolao. The periodic Riccati equation. In Willems Bittanti, Laub, editor, The Riccati Equation. Springer, 1991.

[10] D. Bertsekas. Dynamic Programming and Optimal Control, Athena Scientific Dynamic Programming and Optimal Control, Vol. II: Approximate Dynamic Programming, ISBN-13: 978-1-886529-44-1, 712 pp., hardcover, 2012

[11] L. Chisci and G. Zappa. *Feasibility in predictive control of constrained linear systems: the output feedback case.* International Journal of Robust and Nonlinear Control, 12:465–487, 2002.

[12] M. Diehl and H.J. Ferreau and N. Haverbeke. *Efficient Numerical Methods for Nonlinear MPC and Moving Horizon Estimation*, In Nonlinear Model Predictive Control, 384:391–417, 2009.

[13] M. Diehl, R. Amrit, and J. Rawlings. A Lyapunov Function for Economic Optimizing Model Predictive Control. IEEE Transactions of Automatic Control, 56:703–707, 2011.

[14] H. Chen, F. Allgöwer. A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. Automatica, 34(10):1205–1217, 1998.

[15] A.A. Feldbaum. Dual-control theory (I-IV). Automation and Remote Control, 21, pages 1240–1249 and 1453–1464, 1960, as well 22, pages 3–16 and 129–143, 1961.

[16] N.M. Filatov and H. Unbehauen. Survey of adaptive dual control methods. In Proceedings of the IEEE Conference on Control Theory and Applications, 147(1):118–128, 2000.

[17] N.M. Filatov, H. Unbehauen. Adaptive Dual Control. Springer, 2004.

[18] R. Findeisen, L. Imsland, F. Allgöwer, and B.A. Foss. *State and Output Feedback Nonlinear Model Predictive Control: An Overview*, European Journal of Control, 9:190–207, 2003.

[19] R. Fisher The design of experiments. Oliver & Boyd, 1935.

[20] M. Forgione, X. Bombois, P.M.J. Van den Hof. Data-driven model improvement for model-based control. Automatica, 52:118–124, 2015.

[21] P.J. Goulart and E.C. Kerrigan *Output feedback receding horizon control of constrained systems.* International Journal of Control, 80(1):8-20, 2007.

[22] L. Grüne. Economic receding horizon control without terminal constraints. Automatica 43:725-734, 2011.

[23] N. Haverbeke. Efficient Numerical Methods for Moving Horizon Estimation. Ph.D. Thesis, KU Leuven, 2011. ISBN 978-94-6018-417-8.

[24] T.A.N. Heirung, B.E. Ydstie, B. Foss. Towards Dual MPC. In Proceedings of the 4th IFAC Nonlinear Model Predictive Control Conference, pp:502–507, Noordwijkerhout (Netherlands), 2012.

[25] T.A.N. Heirung, B. Foss, B.E. Ydstie. MPC-based dual control with online experiment design. Journal of Process Control, 32, 64–76, 2015.

[26] T.A.N. Heirung, B.E. Ydstie, B. Foss. Dual MPC for FIR Systems: Information Anticipation. In proceedings of Advanced Control of Chemical Processes, 1034–1039, Whistler, Canada.

[27] B. Hernandez, P. Trodden. Persistently exciting tube MPC. arXiv:1505.05772v1, 2015.

[28] B. Houska, H.J. Ferreau, and M. Diehl. An auto-generated real-time iteration algorithm for nonlinear MPC in the microsecond range. Automatica, 47(10):2279-2285, 2011.

[29] B. Houska, D. Telen, F. Logist, M. Diehl and J. Van Impe. An Economic Objective for Optimal Experiment Design of Nonlinear Dynamic Processes. *Automatica*, 51, pp:98–103, 2015.

[30] B. Houska. Enforcing Asymptotic Orbital Stability of Economic Model Predictive Control. *Automatica*, 57, pp:45–50, 2015.

[31] M. Hovd, R.R. Bitmead. Interaction between control and state estimation in nonlinear MPC. Modeling, Identification, and Control, 26(3):165–174, 2005.

[32] T.K. Kim, T. Sugie. Adaptive receding horizon predictive control for constrained discretetime systems with parameter uncertainties. International Journal of Control, 81(1):62–73, 2008.

[33] A.B. Kurzhanski. *The Problem of Measurement Feedback Control.* Applied Mathematics and Mechanics, 64:547–563, 2004.

[34] C.A. Larsson, M. Annergren, H. Hjalmarsson, C.R. Rojas, X. Bombois, A. Mesbah, P. Moden. Model predictive control with integrated experiment design for output error systems. In proceedings of the 2013 European Control Conference (ECC), 3790–3795, July, 2013.

[35] C. Larsson, C. Rojas, X. Bombois, H. Hjalmarsson. Experimental evaluation of model predictive control with excitation (MPC-X) on an industrial depropanizer. Journal of Process Control, 31:1–16, 2015.

[36] Y.I. Lee and B. Kouvaritakis. Receding-horizon output feedback control for systems with input saturation. IEEE Proceedings on Control Theory and Applications, 148:109–115, 2001.

[37] 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.

[38] S. Lucia, R. Paulen. Robust nonlinear model predictive control with reduction of uncertainty via robust optimal experiment design. In proceedings of the IFAC World Congress, 1904–1909, Cape Town, South Africa, 2014.

[39] G. Marafioti, R.R. Bitmead, M. Hovd. Persistently exciting model predictive control. International Journal of Adaptive Control and Signal Processing, 28(6):536–552, 2014.

[40] D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. Constrained model predictive control: Stability and optimality. Automatica, 36(6):789–814, 2000.

[41] D.Q. Mayne, S.V. Rakovic, R. Findeisen, and F. Allgöwer. Robust output feedback model predictive control of constrained linear systems. *Automatica*, 42:1217–1222, 2006.

[42] D.Q. Mayne, S.V. Rakovic, R. Findeisen, and F. Allgöwer. Robust output feedback model predictive control of constrained linear systems: Time varying case. *Automatica*, 45:2082–2087, 2009.

[43] A. Mesbah, X. Bombois, M. Forgione, H. Hjalmarsson, P.M.J. Van den Hof. Least costly losed-loop performance diagnosis and plant re-identification. International Journal of Control, 88(11):2264–2276, 2015.

[44] J. Nocedal and S.J. Wright. *Numerical Optimization.* Springer Series in Operations Research and Financial Engineering. Springer, 2 edition, 2006.

[45] G.W. Oehlert. A Note on the Delta Method. The American Statistician, 46:27–29, 1992.

[46] F. Pukelsheim. Optimal design of Experiments. John Wiley & Sons, Inc., New York, 1993.

[47] S.V. Rakovic, B. Kouvaritakis, M. Cannon, C. Panos, and R. Findeisen *Parameterized tube model predictive control.* IEEE Transactions on Automatic Control, 57(11), 2012.

[48] C.V. Rao, J.B. Rawlings, and J.H. Lee. *Constrained linear state estimation - a moving horizon approach* Automatica, 37:1619–1628, 2001.

[49] J.B. Rawlings and D.Q. Mayne. Model Predictive Control: Theory and Design. Madison, WI: Nob Hill Publishing, 2009.

[50] S. Sager. Sampling decisions in optimum experimental design in the light of Pontryagin's maximum principle. SIAM Journal on Control and Optimization 51(4), 3181-3207, 2013.

[51] M.S. Shouche, H. Genceli, M. Nikolaou. Effect of online optimization techniques on model predictive control and identidification (MPCI). Computers & Chemical Engineering, 26(9):1241–1252, 2002.

[52] R.F. Stengel. Extended Kalman filter Optimal Control and Estimation, Dover Publications, Inc., 1994.

[53] D. Telen, B. Houska, F. Logist, E. Van Derlinden, M. Diehl, J. Van Impe. Optimal experiment design under process noise using Riccati differential equations. Journal of Process Control, 23:613–629, 2013.

[54] D. Telen, F. Logist, R. Quirynen, B. Houska, M. Diehl, J. Van Impe. Optimal experiment design for nonlinear dynamic (bio)chemical systems using sequential semidefinite programming. AIChE Journal, 60(5), 1728-1739, 2014.

[55] B. Wittenmark and C. Elevitch. An adaptive control algorithm with dual features. In IFAC Symposium on Identification and System Parameter Estimation, pp:587–592, 1985.

[56] B. Wittenmark. Adaptive dual control methods: An overview. In 5th IFAC symposium on Adaptive Systems in Control and Signal Processing Budapest, Hungary, 1995.

[57] J. Yan and R.R. Bitmead. Incorporating state estimation into model predictive control and its application to network traffic control. Automatica, 41(4):595–604, 2005.

[58] E. Zacekova, S. Privara, M. Pcolka. Persistent excitation condition within the dual control framework. Journal of Process Control, 23(9):1270–1280, 2013.

[59] M. Zanon, S. Gros, and M. Diehl. A Lyapunov Function for Periodic Economic Optimizing Model Predictive Control. In Proceedings of the 52nd IEEE Conference on Decision and Control, pp: 5107–5112 Florence, Italy, 2013.