Automatica 94 (2018) 315–323

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

# Stochastic output-feedback model predictive control ☆

Martin A. Sehr [a], Robert R. Bitmead [b,*]

[a] Siemens Corporate Technology, Berkeley, CA 94704, USA
[b] Department of Mechanical & Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093-0411, USA

ARTICLE INFO

*Article history:*
Received 2 June 2017
Received in revised form 22 December 2017
Accepted 16 March 2018

*Keywords:*
Stochastic control
Predictive control
Information state
Performance analysis
Dual optimal control

ABSTRACT

A new formulation of Stochastic Model Predictive Output Feedback Control is presented and analyzed as a transposition of Stochastic Optimal Output Feedback Control into a receding horizon setting. This requires lifting the design into a framework involving propagation of the conditional state density, the *information state*, and solution of the Stochastic Dynamic Programming Equation for an optimal feedback policy, both stages of which are computationally challenging in the general, nonlinear setup. The upside is that the clearance of three bottleneck aspects of Model Predictive Control is connate to the optimality: output feedback is incorporated naturally; dual regulation and probing of the control signal is inherent; closed-loop performance relative to infinite-horizon optimal control is guaranteed. While the methods are numerically formidable, our aim is to develop an approach to Stochastic Model Predictive Control with guarantees and, from there, to seek a less onerous approximation. To this end, we discuss in particular the class of Partially Observable Markov Decision Processes, to which our results extend seamlessly, and demonstrate applicability with an example in healthcare decision making, where duality and associated optimality in the control signal are required for satisfactory closed-loop behavior.


## 1. Introduction

Model Predictive Control (MPC), in its original formulation, is a full-state feedback law. This underpins two theoretical limitations of MPC: accommodation of output feedback, and extension to include a cogent robustness theory since the state dimension is fixed. This paper addresses the first question. There have been a number of approaches, mostly hinging on replacement of the measured true state by a state estimate, which is computed via Kalman filtering (Sehr & Bitmead, 2016; Yan & Bitmead, 2005), moving-horizon estimator (Copp & Hespanha, 2014; Sui, Feng, & Hovd, 2008), tube-based minimax estimators (Mayne, Raković, Findeisen, & Allgöwer, 2009), etc. Apart from Copp and Hespanha (2014), these designs, often for linear systems, separate the estimator design from the control design. The control problem may be altered to accommodate the state estimation error by methods such as: constraint tightening (Yan & Bitmead, 2005), chance/probabilistic constraints (Schwarm & Nikolaou, 1999), and so forth.

In this paper, we first consider Stochastic Model Predictive Control (SMPC), formulated as a variant of Stochastic Optimal Output Feedback Control (SOOFC), without regard to computational tractability restrictions. By taking this route, we establish a formulation of SMPC which possesses central features: accommodation of output feedback and duality/probing; examination of the probabilistic requirements of deterministic and probabilistic constraints; guaranteed performance of the SMPC controller applied to the system. Performance bounds are stated in relation to the infinite-horizon optimally controlled closed-loop performance. We then particularize our performance results to the class of Partially Observed Markov Decision Processes (POMDPs), as is discussed explicitly in Sehr and Bitmead (2018). For this special class of systems, application of our results and verification of the underlying assumptions are computationally tractable, as we demonstrate using a numerical example in healthcare decision-making based on Sehr and Bitmead (2017b).

This paper does *not* seek to provide a comprehensive survey of the myriad alternative approaches proposed for SMPC. For that, we recommend the numerous available references such as Goodwin, Kong, Mirzaeva, and Seron (2014), Kouvaritakis and Cannon (2016), Mayne (2014) and Mesbah (2016). Rather, we present a new algorithm for SMPC based on SOOFC and prove, particularly, performance properties relative to optimality. As a by-product, we acquire a natural treatment of output feedback via the Bayesian Filter and of the associated controller duality required to balance probing for observability enhancement and regulation. The price we pay for general nonlinear systems is the suspension of disbelief

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Riccardo Scattolini under the direction of Editor Ian R. Petersen.

* Corresponding author.
*E-mail addresses:* martin.sehr@siemens.com (M.A. Sehr), rbitmead@ucsd.edu (R.R. Bitmead).

https://doi.org/10.1016/j.automatica.2018.04.013
0005-1098/

in computational tractability. However, the approach delineates a target controller with assured properties. Approximating this intractable controller by a more computationally amenable variant, as opposed to identifying soluble but indirect problems without guarantees, holds the prospect of approximately attracting the benefits. Such a strategy, using a particle implementation of the Bayesian filter and scenario methods at the cost of losing duality of the control inputs, is discussed in Sehr and Bitmead (2017a). Alternatively, as suggested in Sehr and Bitmead (2017b), one may approximate the nonlinear SMPC problem by POMDPs and apply the methods of the current paper directly, resulting in optimality and duality on the approximate POMDP system. We do this to demonstrate the feasibility of the methods and to exemplify and quantify the results claimed.

*Comparison with other performance results*

Our work is related to four central papers discussing performance bounds linking the achieved cost of fully nonlinear MPC on the infinite horizon with the cost of infinite-horizon optimal control:

**Grüne & Rantzer** (2008) study the deterministic, full-state feedback situation and provide comparison between the infinite-horizon optimal cost and the achieved infinite-horizon MPC cost. In particular, the achieved MPC cost is bounded in terms of the computed finite-horizon MPC cost.

**Hernández & Lasserre** (1990) consider the stochastic case with full-state feedback and average as well as discounted costs. Their results yield a comparison between the infinite-horizon stochastic optimal cost and the achieved infinite-horizon MPC cost in terms of the unknown true optimal cost.

**Chatterjee & Lygeros** (2015) also treat the stochastic case with full-state feedback and average cost function. They establish and quantify a bound on the expected long-run average MPC performance related to the terminal cost function and its associated monotonicity requirement.

**Riggs & Bitmead** (2012) consider stochastic full-state feedback as an extension to Grüne and Rantzer (2008) via a discounted infinite-horizon cost function. Similarly to Grüne and Rantzer (2008), they establish a performance bound on the achieved infinite-horizon MPC cost in terms of the computed finite-horizon MPC cost.

**The current paper** extends (Grüne & Rantzer, 2008; Riggs & Bitmead, 2012) to include output feedback stochastic MPC. Achieved performance is bounded in terms of the computed finite-horizon MPC cost. The native incorporation of state estimation into the problem is the central contribution.

Each of these works relies on a sequence of assumptions concerning the well-posedness of the underlying optimization problems and specific monotonicity conditions on certain value functions which admit the establishment of stability and performance bounds. We indicate the universality of this class of assumptions in prior performance bounds for MPC.

*Main contributions*

We summarize the main theoretical/technical contribution of this paper, Corollary 2, for stochastic MPC with state estimation. Subject to cost monotonicity Assumption 10, which is testable in terms of a known terminal policy and the terminal cost function, an upper bound is computable for the achieved discounted infinite-horizon SMPC cost in terms of the computed finite-horizon SMPC cost and other parameters of the monotonicity condition. The central practical contribution is partially negative, in tying output feedback performance guarantees to complicated online algorithms involving the propagation of conditional probability density functions, but also positive in highlighting the necessity of duality and probing in order to enhance state observability and in providing a target optimal controller worthy of close approximation. The principle theoretical results of MPC are based on optimal control to yield stability and performance in practice. There is a strong move in the practice of signal processing and control towards increasingly onerous realtime computations, such as Maximum-Likelihood and Particle Filter estimation or scenario-based methods. Likewise, explicit MPC devotes extraordinary computer power to the offline solution of receding horizon optimal policies.

To illustrate the feasibility of this approach to SMPC in a particular subset of problems, we provide an example – here a POMDP from healthcare – in which the assumptions are verifiable and verified, indicating their substance and the nature of the qualitative and quantified conclusions regarding closed-loop output-feedback stochastic MPC. To aid the development, all proofs are relegated to the Appendix. We write sequences as $\mathbf{t}^m \triangleq \{t_0, t_1 \ldots, t_m\}$.

## 2. Stochastic optimal output-feedback control

We consider stochastic optimal control of nonlinear time-invariant dynamics of the form

$$x_{k+1} = f(x_k, u_k, w_k), \quad x_0, \tag{1}$$

$$y_k = h(x_k, v_k), \tag{2}$$

where $k \in \mathbb{N}_0$, $x_k \in \mathbb{R}^{n_x}$ denotes the state with initial value $x_0$, $u_k \in \mathbb{R}^{n_u}$ the control input, $y_k \in \mathbb{R}^{n_y}$ the measurement output, $w_k \in \mathbb{R}^{n_w}$ the process noise and $v_k \in \mathbb{R}^{n_v}$ the measurement noise.

We denote by

$$\pi_{0|-1} \triangleq \text{pdf}(x_0), \tag{3}$$

the known a-priori density of the initial state and by

$$\zeta^k \triangleq \{y_0, u_0, y_1, u_1, \ldots, u_{k-1}, y_k\}, \qquad \zeta^0 \triangleq \{y_0\},$$

the data available at time $k$. We make the following standing assumptions on the random variables and system dynamics.

**Assumption 1.** The dynamics (1)–(2) satisfy

1. $f(\cdot, u, \cdot)$ is differentiable a.e. with full rank Jacobian $\forall\, u \in \mathbb{R}^{n_u}$.
2. $h(\cdot, \cdot)$ is differentiable a.e. with full rank Jacobian.
3. $w_k$ and $v_k$ are i.i.d. sequences with known densities.
4. $x_0, w_k, v_l$ are mutually independent for all $k, l \geq 0$.

**Assumption 2.** The control input $u_k$ at time instant $k \geq 0$ is a function of the data $\zeta^k$ and $\pi_{0|-1}$.

As there is no direct feedthrough from $u_k$ to $y_k$, Assumptions 1 and 2 assure that system (1)–(2) is a *controlled Markov process* (Kumar & Varaiya, 1986). Assumption 1 further ensures that $f$ and $h$ enjoy the *Ponomarev 0-property* (Ponomarev, 1987) and hence that $x_k$ and $y_k$ possess joint and marginal densities possibly together with a finite number of discontinuities, which feature is used in extending the probability density function calculations to probability mass functions for POMDPs in Sections 4 and 5. It precludes the existence of a singular continuous part (Rudin, 1974) of the measures. Satisfaction of Assumption 1 admits the analysis of the stochastic behavior using sampling methods such as the Particle Filter or discrete simulation.