IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 25 September 2023, accepted 16 October 2023, date of publication 20 October 2023, date of current version 26 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3326344

RESEARCH ARTICLE

Probabilistic Multi-Step Identification With Implicit State Estimation for Stochastic MPC

FELIX FIEDLER AND SERGIO LUCIA, (Member, IEEE)
Chair of Process Automation Systems, TU Dortmund University, 44227 Dortmund, Germany
Corresponding author: Felix Fiedler (felix.fiedler@tu-dortmund.de)

ABSTRACT Stochastic Model Predictive Control (SMPC) is a promising solution for controlling multivariable systems in the presence of uncertainty. However, a core challenge lies in obtaining a probabilistic system model. Recently, multi-step system identification has been proposed as a solution. Multi-step models simultaneously predict a finite sequence of future states, which traditionally involves recursive evaluation of a state-space model. Particularly in the stochastic context, the recursive evaluation of identified state-space models has several drawbacks, making multi-step models an appealing choice. As a main novelty of this work, we propose a probabilistic multi-step identification method for a linear system with noisy state measurements and unknown process and measurement noise covariances. We show that, in expectation, evaluating the identified multi-step model is equivalent to estimating the initial state distribution and subsequently propagating this distribution using the known system dynamics. Therefore, using only recorded data of an unknown linear system, our proposed method yields a probabilistic multi-step model, including the state estimation task, that can be directly used for SMPC. As an additional novelty, our proposed SMPC formulation considers parametric uncertainties of the identified multi-step model. We demonstrate our method in two simulation studies, showcasing its effectiveness even for a nonlinear system with output feedback.

INDEX TERMS Stochastic model predictive control, system identification, multi-step identification, data-based control.

I. INTRODUCTION

Model predictive control (MPC) is a popular strategy to control multivariable systems with constraints [1]. At its core, MPC uses a system model to predict and optimize the future behavior of the system. Obtaining a suitable system model is therefore a central requirement for MPC and has been recognized as a major challenge for decades [2]. Traditionally, state-space representations of the prediction model have been used for MPC and other control schemes. A state-space representation is the obvious choice for models that are obtained from first principles and typically employed for data-based system identification [3], [4]. Under ideal conditions, state-space models can exactly describe the system dynamics with the fewest number of parameters. To obtain predictions over multiple timesteps, as required for MPC, the state-space model can be evaluated recursively. However, a recent trend in system identification is to directly identify multi-step prediction models [5], [6], [7], [8].

A multi-step model can simultaneously predict finite sequences of future states with an individual function for each step of the sequence. At first glance, a multi-step model appears to add unnecessary complexity to the system identification and control task. However, recent work has shown that multi-step identification can have significant advantages over state-space identification [7], [8]. Most importantly, multi-step models can have a better accuracy than state-space models and complexity is added only to the offline identification task, not to the online control task.

A closely related trend to multi-step model identification is data-enabled predictive control (DeePC) [9], [10], [11], [12]. DeePC draws from behavioral systems theory and combines an implicit multi-step model identification and control task in a single optimization problem. The close relationship

The associate editor coordinating the review of this manuscript and approving it for publication was Ton Duc Do.

between DeePC and MPC with multi-step prediction models has recently been shown in [12], [13].

Another major challenge for the application of MPC is the presence of uncertainty. While nominal MPC provides a certain robustness, it is often not sufficient, especially if the system operates close to safety-critical constraints. A promising approach to handle this situation is stochastic MPC (SMPC) [14], [15]. SMPC can be applied when the uncertainty in the model follows a known probability distribution function and it relies on the formulation of chance constraints, that is, constraints with specified probability of violation.

Tackling the challenges of data-based identification and control under uncertainty with state-space and multi-step models is a prominent field of research [6], [7], [8], [17]. There are, however, important limitations in previous works using multi-step identification and SMPC. For example, it was previously assumed that the process and measurement noise covariances are known for the identification of the probabilistic multi-step model [8]. Unfortunately, this assumption significantly limits the applicability in practice, as the process noise typically represents unknown disturbances of the system. Additionally, most stochastic MPC formulations require state-feedback to solve the optimal control problem [14], [15], which is often not available in practice.

As a main contribution of this work, we propose a stochastic MPC formulation based on multi-step identification without knowledge of the process and measurement noise covariances. Our proposed method is derived for systems with noisy state measurements and we formulate the objective function and chance constraints of the SMPC problem in terms of these measurements. We show that with maximum likelihood estimation, we identify a multi-step model that, in expectation, describes the true distribution of the future measurements of the system. Crucially, this true distribution also considers the uncertainty for the initial state which is estimated from the initial measurement using a Kalman filter [1]. Therefore, evaluating our identified multi-step model with noisy state measurements is equivalent to estimating the initial state distribution and subsequently propagating this distribution using the known system dynamics. Finally, we show that this property is a unique advantage of multi-step models and does not apply to recursively evaluated state-space models.

The identified multi-step model is thus readily applicable for the formulation of an SMPC problem with noisy state-feedback. As another contribution of this work, our proposed SMPC formulation directly considers the parametric uncertainties of the identified multi-step model. This is in contrast to previous work, where the parametric uncertainties were included as an ellipsoidal uncertainty set [8]. As a final contribution, we investigate the proposed SMPC controller based on multi-step models, comparing it to a variant relying on identified state-space models. We conduct this evaluation in two simulation studies, showcasing the effectiveness of SMPC with identified multi-step model, even for a nonlinear system with output feedback.

This paper is structured as follows. In Section II, we introduce the stochastic optimal control problem for a linear dynamic system with known system matrices and noise covariances. In Section III, we assume that this information is not available and discuss probabilistic multi-step identification from data. In Section IV, the proposed SMPC problem based on the identified multi-step model is presented. We investigate our method with a linear building system in Section VI, and a nonlinear CSTR system in Section VII. The work is concluded in Section VIII.

II. STOCHASTIC MODEL PREDICTIVE CONTROL

We consider a linear and uncertain dynamic system given in the discrete-time formulation as:

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{E}_x\boldsymbol{e}_{x,k}, \tag{1a}$$

$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{E}_y\boldsymbol{e}_{y,k}, \tag{1b}$$

with states $\boldsymbol{x} \in \mathbb{R}^{n_x}$, inputs $\boldsymbol{u} \in \mathbb{R}^{n_u}$, measurements $\boldsymbol{y} \in \mathbb{R}^{n_y}$, process noise $\boldsymbol{e}_{x,k} \in \mathbb{R}^{n_{e,x}}$ and measurement noise $\boldsymbol{e}_{y,k} \in \mathbb{R}^{n_{e,y}}$. The system is described with the matrices $\boldsymbol{A} \in \mathbb{R}^{n_x \times n_x}$, $\boldsymbol{B} \in \mathbb{R}^{n_x \times n_u}$, $\boldsymbol{C} \in \mathbb{R}^{n_y \times n_x}$, $\boldsymbol{E}_x \in \mathbb{R}^{n_x \times n_{e,x}}$ and $\boldsymbol{E}_y \in \mathbb{R}^{n_y \times n_{e,y}}$. As in previous related work [8], we introduce the following assumption.

Assumption 1: The system (1) is subject to noisy state-feedback, that is $\boldsymbol{C} = \boldsymbol{I}$.

For practical applications, this assumption can be relaxed to output-feedback, as we demonstrate in Section VII. Furthermore, we assume the following properties of the system noise.

Assumption 2: The system (1) is subject to additive Gaussian process noise and measurement noise, that is: $\boldsymbol{e}_{x,k} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_x)$ and $\boldsymbol{e}_{y,k} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_y)$, $\forall k$.

As a main premise of this work, we consider the system matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{E}_x$ and $\boldsymbol{E}_y$ and the covariance matrices $\boldsymbol{\Sigma}_x$ and $\boldsymbol{\Sigma}_y$ as unknown. In the following subsection, we derive multi-step models from the state-space representation in (1), and propose an identification method of the parameters of the multi-step model in Section III.

A. MULTI-STEP PREDICTIONS

To formulate the stochastic model predictive control problem in Subsection II-C, we require the probability distribution of the future sequence of system measurements. If the system matrices are known, a multi-step prediction can be obtained by recursively evaluating the state state-space model in (1). Considering Assumption 1, we have:

$$\boldsymbol{y}_{[1,N]} = \mathcal{O}(\boldsymbol{A})\boldsymbol{x}_0 + \mathcal{T}(\boldsymbol{A}, \boldsymbol{B})\boldsymbol{u}_{[0,N-1]} + \mathcal{T}(\boldsymbol{A}, \boldsymbol{E}_x)\boldsymbol{e}_{x,[0,N-1]} + (\boldsymbol{I}_N \otimes \boldsymbol{E}_y)\boldsymbol{e}_{y,[1,N]}, \tag{2}$$

where $\boldsymbol{y}_{[1,N]}^\top = \left[\boldsymbol{y}_1^\top, \ldots, \boldsymbol{y}_N^\top\right]$ denotes a finite sequence of outputs, $\otimes$ is the Kronecker product and $\boldsymbol{I}_N$ is the identity matrix of dimension N. The element in block i, j of the matrices $\mathcal{T}(\cdot,\cdot)$ and $\mathcal{O}(\cdot)$ can be obtained as:

$$[\mathcal{T}(\boldsymbol{A},\boldsymbol{B})]_{i,j} = \begin{cases} \boldsymbol{A}^{i-j}\boldsymbol{B}, & \text{if } i \geq j \\ \boldsymbol{0}, & \text{otherwise} \end{cases} \quad \forall i,j \in \mathbb{I}_{[1,N]}, \tag{3a}$$

$$[\mathcal{O}(\boldsymbol{A})]_{i,1} = \boldsymbol{A}^{i} \quad \forall i \in \mathbb{I}_{[1,N]}. \tag{3b}$$

We denote with $\mathbb{I}_{[1,N]}$ the set of integers from 1 to N. Considering (2) and Assumption 2, it follows for the distributed multi-step predictions:

$$p(\boldsymbol{y}_{[1,N]}|\boldsymbol{x}_0, \boldsymbol{u}_{[0,N-1]}) = \mathcal{N}(\bar{\boldsymbol{y}}_{[1,N]}, \boldsymbol{\Sigma}_{y,[1,N]}), \tag{4a}$$

with:

$$\bar{\boldsymbol{y}}_{[1,N]} = \mathcal{O}(\boldsymbol{A})\boldsymbol{x}_0 + \mathcal{T}(\boldsymbol{A},\boldsymbol{B})\boldsymbol{u}_{[0,N-1]}, \tag{4b}$$

$$\boldsymbol{\Sigma}_{y,[1,N]} = \mathcal{T}(\boldsymbol{A},\boldsymbol{E}_x)(\boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_x)\mathcal{T}(\boldsymbol{A},\boldsymbol{E}_x)^\top + \boldsymbol{I}_N \otimes (\boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top). \tag{4c}$$

The covariance in (4c) is obtained by considering the distributions $\boldsymbol{e}_{x,[0,N-1]} \sim \mathcal{N}(0, \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_x)$ and $\boldsymbol{e}_{y,[1,N]} \sim \mathcal{N}(0, \boldsymbol{I}_N \otimes \boldsymbol{\Sigma}_y)$, the properties of a linear transformation of a normally distributed variable [18, 20.23 b], and the mixed product rule for the Kronecker product [18, 11.11 a].

B. INCORPORATING THE STATE ESTIMATION

The predictive distribution in (4) is conditional on the initial state $\boldsymbol{x}_0$. However, due to the presence of measurement noise, the initial state $\boldsymbol{x}_0$ is unknown and must be estimated. The optimal state estimator for system (1) and considering Assumption 2 is the *Kalman filter* [1]. The Kalman filter yields a distribution for the estimated initial state which should also be considered for the distributed multi-step prediction.

In the following lemma, we incorporate the estimated distribution of the initial state to obtain the distribution $p(\hat{\boldsymbol{y}}_{[1,N]}|\boldsymbol{y}_0, \boldsymbol{u}_{[0,N-1]})$. In contrast to (4), this distribution is conditional on the noise affected measurement $\boldsymbol{y}_0$ instead of the unknown $\boldsymbol{x}_0$. Obtaining this distribution allows to directly state the stochastic optimal control problem for the measured initial state in the following subsection. Furthermore, we will revisit this distribution for the multi-step identification in Section III.

For the state estimation, we require a prior distribution of $\boldsymbol{x}_0$. Without loss of generality and for ease of notation, we assume that this prior has a zero mean.

Assumption 3: The states of the dynamic system are distributed according to $\boldsymbol{x}_0 \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{x,0})$.

This allows to state the following lemma.

Lemma 1: Assumption 1-3 hold. The distribution of the multi-step prediction, conditional on $\boldsymbol{y}_0$ *and* $\boldsymbol{u}_{[0,N-1]}$*, can be described with*:

$$p(\hat{\boldsymbol{y}}_{[1,N]}|\boldsymbol{y}_0, \boldsymbol{u}_{[0,N-1]}) = \mathcal{N}(\bar{\hat{\boldsymbol{y}}}_{[1,N]},\ \hat{\boldsymbol{\Sigma}}_{y,[1,N]}). \tag{5}$$

The mean and covariance in (5) are:

$$\bar{\hat{\boldsymbol{y}}}_{[1,N]} = \mathcal{O}(\boldsymbol{A})\boldsymbol{L}\boldsymbol{y}_0 + \mathcal{T}(\boldsymbol{A},\boldsymbol{B})\boldsymbol{u}_{[0,N-1]}, \tag{6a}$$

$$\hat{\boldsymbol{\Sigma}}_{y,[1,N]} = \boldsymbol{\Sigma}_{y,[1,N]} + \mathcal{O}(\boldsymbol{A})(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{\Sigma}_{x,0}\mathcal{O}(\boldsymbol{A})^\top, \tag{6b}$$

where $\boldsymbol{\Sigma}_{y,[1,N]}$ *is defined in (4c), and the Kalman gain can be computed as:*

$$\boldsymbol{L} = \boldsymbol{\Sigma}_{x,0}(\boldsymbol{\Sigma}_{x,0} + \boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top)^{-1}. \tag{6c}$$

Proof: With observed measurements $\boldsymbol{y}_0$*, and prior distribution of the states* $\boldsymbol{x}_0$ *from Assumption 3, we can perform a Kalman filter correction step [1]:*

$$\begin{aligned} \boldsymbol{L} &= \boldsymbol{\Sigma}_{x,0}\boldsymbol{C}^\top(\boldsymbol{C}\boldsymbol{\Sigma}_{x,0} + \boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top)^{-1}, \\ \bar{\boldsymbol{x}}_0^+ &= \bar{\boldsymbol{x}}_0 + \boldsymbol{L}(\boldsymbol{y}_0 - \boldsymbol{C}\bar{\boldsymbol{x}}_0), \\ \boldsymbol{\Sigma}_{x,0}^+ &= (\boldsymbol{I} - \boldsymbol{L}\boldsymbol{C})\boldsymbol{\Sigma}_{x,0}, \end{aligned}$$

with prior mean $\bar{\boldsymbol{x}}_0 = 0$*, posterior mean* $\bar{\boldsymbol{x}}_0^+$ *and posterior covariance* $\boldsymbol{\Sigma}_{x,0}^+$*. With* $\boldsymbol{C} = \boldsymbol{I}$*, due to Assumption 1, and* $\bar{\boldsymbol{x}}_0 = 0$*, due to Assumption 3, we obtain:*

$$\begin{aligned} \boldsymbol{L} &= \boldsymbol{\Sigma}_{x,0}(\boldsymbol{\Sigma}_{x,0} + \boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top)^{-1}, \\ \bar{\boldsymbol{x}}_0^+ &= \boldsymbol{L}\boldsymbol{y}_0, \\ \boldsymbol{\Sigma}_{x,0}^+ &= (\boldsymbol{I} - \boldsymbol{L})\boldsymbol{\Sigma}_{x,0}, \end{aligned}$$

With mean $\bar{\boldsymbol{x}}_0^+$ *and covariance* $\boldsymbol{\Sigma}_{x,0}^+$*, we have the posterior distribution:*

$$\boldsymbol{x}_0 \sim \mathcal{N}(\boldsymbol{L}\boldsymbol{y}_0,\ (\boldsymbol{I}-\boldsymbol{L})\boldsymbol{\Sigma}_{x,0}). \tag{7}$$

Substituting the distribution (7) in (2) yields (6). ☐

We omit the prediction step of the Kalman filter in Lemma 1 for reasons that will be clarified in Subsection IV-A. For the discussion in this section, the prediction step is implicitly considered through Assumption 3, that is, we assume to have the true prior distribution of the states.

C. STOCHASTIC OPTIMAL CONTROL PROBLEM

With the distributed multi-step system response (5), we can now state the stochastic optimal control problem. As is commonly the case for data-based MPC [9], we formulate cost and constraints for the measured outputs of the system. Furthermore, we directly incorporate the state estimation in the formulation and consider the uncertainty of the distributed initial state as shown in the previous subsection.

With expectation $\mathbb{E}(\cdot)$, and probability $P(\cdot)$, we have the stochastic optimal control problem:

$$\begin{aligned} \min_{\boldsymbol{u}_{[0,N-1]}\in\mathbb{A}} \quad & \mathbb{E}\left[\|\hat{\boldsymbol{y}}_{[1,N]}\|_{\boldsymbol{Q}}^2 + \|\boldsymbol{u}_{[0,N-1]}\|_{\boldsymbol{R}}^2\right] \\ \text{s.t.} \quad & P\left[\boldsymbol{a}_j^\top\hat{\boldsymbol{y}}_{[1,N]} \leq b_j\right] \geq (1-\epsilon) \quad \forall j \in \mathbb{I}_{[1,n_c]}, \end{aligned} \tag{8}$$

with $\hat{\boldsymbol{y}}_{[1,N]} \sim \mathcal{N}(\bar{\hat{\boldsymbol{y}}}_{[1,N]}, \hat{\boldsymbol{\Sigma}}_{y,[1,N]})$ obtained from (5), positive definite weighting matrices $\boldsymbol{Q} \succ 0$, $\boldsymbol{R} \succ 0$ and convex constraint set $\mathbb{A}$. We denote the norm $\|\boldsymbol{y}\|_{\boldsymbol{Q}}^2 = \boldsymbol{y}^\top\boldsymbol{Q}\boldsymbol{y}$. The n_c *chance constraints*, can be violated with probability $\epsilon \in]0, 1[$, and are introduced as individual halfspaces [19, 2.2.1] with $\boldsymbol{a}_j \in \mathbb{R}^{n_y}$ and $b_j \in \mathbb{R}$ for $j \in \mathbb{I}_{[1,n_c]}$.

In the described setting, we can reformulate (8) as a deterministic problem, which yields the same optimal solution [15]:

$$\begin{aligned}\min_{\boldsymbol{u}_{[0,N-1]}\in\mathbb{A}} \ & \|\hat{\bar{\boldsymbol{y}}}_{[1,N]}\|_{\boldsymbol{Q}}^2 + \|\boldsymbol{u}_{[0,N-1]}\|_{\boldsymbol{R}}^2 + \operatorname{trace}\left(\boldsymbol{Q}\hat{\boldsymbol{\Sigma}}_{y,[1,N]}\right) \\ \text{s.t.} \ & \boldsymbol{a}_j^\top \hat{\bar{\boldsymbol{y}}}_{[1,N]} \le b_j - c_p(\epsilon)\|\boldsymbol{a}_j\|_{\hat{\boldsymbol{\Sigma}}_{y,[1,N]}} \quad \forall j \in \mathbb{I}_{[1,n_c]},\end{aligned} \tag{9}$$

with $\hat{\bar{\boldsymbol{y}}}_{[1,N]}$ and $\hat{\boldsymbol{\Sigma}}_{y,[1,N]}$ from (6). The factor $c_p(\epsilon) = \sqrt{2}\operatorname{erf}^{-1}(1-2\epsilon)$ is the quantile of the standard normal distribution and erf^{-1} denotes the inverse error-function [20].

Problem (9) is a quadratic optimization problem depending on the measured initial state $\boldsymbol{y}_0$ and implicitly considers a Kalman filter for the state estimation.

III. PROBABILISTIC MULTI-STEP IDENTIFICATION

For the formulation of the stochastic optimal control problem (9), we require the multi-step distribution in (5). Commonly, the parameters of this distribution are obtained from the system matrices in (1) and the covariances of process and measurement noise from Assumption 2. However, In many practical applications this information is unavailable.

In this section, we propose to use maximum likelihood estimation to identify directly the parameters of the multi-step distribution in (5) from system data. In contrast to previous work [8], we do not assume knowledge of the process and measurement noise covariances. As a main contribution, we present Theorem 1. The theorem establishes that, with the expected values for the estimated parameters and covariance matrix, we obtain the distributed multi-step prediction in (5). The identified multi-step model thus implicitly estimates the initial state distribution from noisy measurements.

A. PRELIMINARIES

To simplify the successive notation, we formulate the multi-step prediction model (2) as:

$$\boldsymbol{t} = \boldsymbol{W}^\top \boldsymbol{z} + \boldsymbol{e}_t, \tag{10}$$

with independent variable $\boldsymbol{z} = [\boldsymbol{x}_0^\top, \ \boldsymbol{u}_{[0,N-1]}^\top]^\top \in \mathbb{R}^{n_z}$, and response variable $\boldsymbol{t} = \boldsymbol{y}_{[1,N]} \in \mathbb{R}^{n_t}$. We have $n_z = n_x + Nn_u$ independent variables and $n_t = Nn_y$ response variables. The transposed parameters are

$$\boldsymbol{W}^T = [\mathcal{O}(\boldsymbol{A}), \ \mathcal{T}(\boldsymbol{A}, \boldsymbol{B})], \tag{11}$$

and we have additive noise $\boldsymbol{e}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_t)$, with $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_{y,[1,N]}$. It follows directly for the distribution of the response variables:

$$p(\boldsymbol{t}|\boldsymbol{z}, \boldsymbol{W}, \boldsymbol{\Sigma}_t) = \mathcal{N}(\boldsymbol{W}^\top \boldsymbol{z}, \ \boldsymbol{\Sigma}_t). \tag{12}$$

Due to measurement noise, the true value of $\boldsymbol{z}$ is unknown and we introduce

$$\boldsymbol{v} = \boldsymbol{z} + \boldsymbol{e}_v, \tag{13}$$

with $\boldsymbol{e}_v \sim \mathcal{N}(0, \boldsymbol{\Sigma}_v)$ and $\boldsymbol{\Sigma}_v = \operatorname{diag}(\boldsymbol{E}_y \boldsymbol{\Sigma}_y \boldsymbol{E}_y^\top, \boldsymbol{0})$. We denote with $\operatorname{diag}(\boldsymbol{A}, \boldsymbol{B})$ the block-diagonal matrix with $\boldsymbol{A}$ and $\boldsymbol{B}$ on the diagonal.

B. PARAMETER AND COVARIANCE ESTIMATION

We seek to identify the multi-step system response from data. To this end, we gather m sequences of the dynamic system and introduce:

$$\boldsymbol{v}_i = \left[\boldsymbol{y}_0^{(i)\top}, \ \boldsymbol{u}_{[0,N-1]}^{(i)\top}\right]^\top, \ \boldsymbol{t}_i = \boldsymbol{y}_{[1,N]}^{(i)} \quad \forall i \in \mathbb{I}_{[1,m]}, \tag{14}$$

We have the set $\mathcal{D} = \{\boldsymbol{V}, \boldsymbol{T}\}$ with *design matrix* [21] $\boldsymbol{V} = [\boldsymbol{v}_i, \ldots, \boldsymbol{v}_m]^\top$ and $\boldsymbol{T} = [\boldsymbol{t}_i, \ldots, \boldsymbol{t}_m]^\top$. The linear model (10) is compactly evaluated for all samples with:

$$\boldsymbol{T} = \boldsymbol{V}\boldsymbol{W} + \boldsymbol{E}_t, \tag{15}$$

where the vectorized residual matrix $\boldsymbol{E}_t$ has the distribution $\operatorname{vec}(\boldsymbol{E}_t) \sim \mathcal{N}(0, \boldsymbol{\Sigma}_t \otimes \boldsymbol{I}_m)$. We require the following assumption for the design matrix.

Assumption 4: We have $m > n_z$ samples and the design matrix $\boldsymbol{V} \in \mathbb{R}^{m \times n_z}$ has full rank, that is, rank$(\boldsymbol{V}) = n_z$.

In practice, Assumption 4 also implies that the gathered input sequences must be persistently exciting [10].

This allows to state the following theorem for the identification of the probabilistic multi-step model.

Theorem 1: Assumption 1 and 2 hold. We have data samples from (14) that satisfy Assumption 3 and 4. The bias corrected maximum likelihood approach yields the estimated parameters and covariance matrix:

$$\hat{\boldsymbol{W}}^* = \boldsymbol{\Sigma}_p^* \boldsymbol{V}^\top \boldsymbol{T}, \tag{16a}$$

$$\hat{\boldsymbol{\Sigma}}_t^* = \frac{1}{m - n_z}(\boldsymbol{V}\hat{\boldsymbol{W}}^* - \boldsymbol{T})^\top(\boldsymbol{V}\hat{\boldsymbol{W}}^* - \boldsymbol{T}), \tag{16b}$$

with:

$$\hat{\boldsymbol{\Sigma}}_p^* = (\boldsymbol{V}^\top \boldsymbol{V})^{-1}. \tag{16c}$$

We partition the estimated parameters as

$$\hat{\boldsymbol{W}}^{*\top} = [\hat{\mathcal{O}}_{\boldsymbol{A}}^*, \ \hat{\mathcal{T}}_{\boldsymbol{A},\boldsymbol{B}}^*],$$

and have $\hat{\boldsymbol{\Sigma}}_t^* = \hat{\boldsymbol{\Sigma}}_{y,[1,N]}^*$. The estimated parameters and covariance matrix have the property:

$$\mathbb{E}[\hat{\mathcal{O}}_{\boldsymbol{A}}^*] = \mathcal{O}(\boldsymbol{A})\boldsymbol{L}, \tag{17a}$$

$$\mathbb{E}[\hat{\mathcal{T}}_{\boldsymbol{A},\boldsymbol{B}}^*] = \mathcal{T}(\boldsymbol{A}, \boldsymbol{B}), \tag{17b}$$

$$\mathbb{E}[\hat{\boldsymbol{\Sigma}}_{y,[1,N]}^*] = \boldsymbol{\Sigma}_{y,[1,N]} + \mathcal{O}(\boldsymbol{A})(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{\Sigma}_{x,0}\mathcal{O}(\boldsymbol{A})^\top. \tag{17c}$$

In expectation, the estimated parameters and covariance matrix are thus identical to the true parameters and covariance matrix in (5).

Proof: We consider the multi-step model in the form of (12) and have noise disturbed independent variables from (13). Only as an intermediate step of the derivation, we introduce the assumption that $\boldsymbol{u}_{[0,N-1]} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{u,0})$. This allows to state the distribution $\boldsymbol{z} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_z)$ with

$\boldsymbol{\Sigma}_z = diag(\boldsymbol{\Sigma}_{x,0}, \boldsymbol{\Sigma}_{u,0})$. *We state the joint distribution of* $\boldsymbol{t}$ *and* $\boldsymbol{v}$ *as shown in [22]:*

$$p\left([\mathbf{t}, \mathbf{v}]^\top | \boldsymbol{W}, \boldsymbol{\Sigma}_t\right) = \mathcal{N}\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\boldsymbol{W}^\top\boldsymbol{\Sigma}_z\boldsymbol{W} + \boldsymbol{\Sigma}_t & \boldsymbol{W}^\top\boldsymbol{\Sigma}_z \\ \boldsymbol{\Sigma}_z\boldsymbol{W} & \boldsymbol{\Sigma}_z + \boldsymbol{\Sigma}_v\end{bmatrix}\right), \tag{18}$$

and the resulting conditional probability follows directly from the properties of a Gaussian normal distribution:

$$p(\mathbf{t}|\mathbf{v}, \boldsymbol{W}, \boldsymbol{\Sigma}_t) = \mathcal{N}(\boldsymbol{W}^\top\boldsymbol{K}\boldsymbol{v},\ \boldsymbol{\Sigma}_t + \boldsymbol{W}^\top(\boldsymbol{\Sigma}_z - \boldsymbol{K}\boldsymbol{\Sigma}_z)\boldsymbol{W}), \tag{19}$$

with reliability matrix [23]:

$$\boldsymbol{K} = \boldsymbol{\Sigma}_z(\boldsymbol{\Sigma}_z + \boldsymbol{\Sigma}_v)^{-1}. \tag{20}$$

Furthermore, we introduce $\hat{\boldsymbol{W}}^\top = \boldsymbol{W}^\top\boldsymbol{K}$ *and* $\hat{\boldsymbol{\Sigma}}_t = \boldsymbol{\Sigma}_t + \boldsymbol{W}^\top(\boldsymbol{\Sigma}_z - \boldsymbol{K}\boldsymbol{\Sigma}_z)\boldsymbol{W}$ *and state (19) as:*

$$p(\mathbf{t}|\mathbf{v}, \hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_t) = \mathcal{N}(\hat{\boldsymbol{W}}^\top\mathbf{v},\ \hat{\boldsymbol{\Sigma}}_t). \tag{21}$$

Of the independent variables, only $\mathbf{y}_0$ *is affected by measurement error. We therefore introduce* $\boldsymbol{\Sigma}_v = diag(\boldsymbol{E}\boldsymbol{\Sigma}_y\boldsymbol{E}^\top, \mathbf{0})$*, and can further simplify (20):*

$$\boldsymbol{K} = \begin{bmatrix}\boldsymbol{\Sigma}_{x,0} & 0 \\ 0 & \boldsymbol{\Sigma}_{u,0}\end{bmatrix}\left(\begin{bmatrix}\boldsymbol{\Sigma}_{x,0} + \boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top & 0 \\ 0 & \boldsymbol{\Sigma}_{u,0}\end{bmatrix}\right)^{-1} = \begin{bmatrix}\boldsymbol{\Sigma}_{x,0}(\boldsymbol{\Sigma}_{x,0} + \boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top)^{-1} & 0 \\ 0 & \boldsymbol{I}\end{bmatrix} \overset{(6c)}{=} \begin{bmatrix}\boldsymbol{L} & 0 \\ 0 & \boldsymbol{I}\end{bmatrix}.$$

The previously introduced $\boldsymbol{\Sigma}_{u,0}$ *thus vanishes from the expression. We substitute* $\boldsymbol{K} = diag(\boldsymbol{L}, \boldsymbol{I})$ *in (19) and consider the definition of the parameters* $\boldsymbol{W}$ *from (11). We obtain:*

$$\hat{\boldsymbol{W}}^\top = [\mathcal{O}(\boldsymbol{A})\boldsymbol{L},\ \mathcal{T}(\boldsymbol{A}, \boldsymbol{B})], \tag{22a}$$

$$\hat{\boldsymbol{\Sigma}}_t = \boldsymbol{\Sigma}_{y,[1,N]} + \mathcal{O}(\boldsymbol{A})(\boldsymbol{I} - \boldsymbol{L})\boldsymbol{\Sigma}_{x,0}\mathcal{O}(\boldsymbol{A})^\top. \tag{22b}$$

With the distribution of the response variables in (21), we have the likelihood:

$$p(\boldsymbol{t} = \boldsymbol{T}|\boldsymbol{V}, \hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_t) = p(\mathcal{D}|\hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_t). \tag{23}$$

To compute the likelihood, we vectorize (15) and consider the properties of the Kronecker product [18, 11.16]:

$$p(\mathcal{D}|\hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_t) = \mathcal{N}((\boldsymbol{I} \otimes \boldsymbol{V})vec(\hat{\boldsymbol{W}}),\ \hat{\boldsymbol{\Sigma}}_t \otimes \boldsymbol{I}_m). \tag{24}$$

We then maximize the likelihood:

$$\hat{\boldsymbol{W}}^*,\ \hat{\boldsymbol{\Sigma}}_t^* = \underset{\hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_t}{\arg\max}\ p(\mathcal{D}|\hat{\boldsymbol{W}}, \hat{\boldsymbol{\Sigma}}_t). \tag{25}$$

Problem (25) has the explicit solution shown in (16) with bias corrected covariance matrix [21], [23]. The inverse in (16c) always exists and we have $m > n_z$ *in (16b) due to Assumption 4. The estimated parameters and covariance have the property that [21], [23]:*

$$\mathbb{E}(\hat{\boldsymbol{W}}^*) = \hat{\boldsymbol{W}},$$
$$\mathbb{E}(\hat{\boldsymbol{\Sigma}}_t^*) = \hat{\boldsymbol{\Sigma}}_t.$$

Considering the definition of $\hat{\boldsymbol{W}}$ *and* $\hat{\boldsymbol{\Sigma}}_t$ *in (22), we obtain (17).* □

As a main consequence of Theorem 1, we have that the estimated parameters and covariance in (16) are, in expectation, identical to the parameters and covariance of the distribution (5). Therefore, we directly obtain the distributed multi-step predictions that incorporates the state estimation.

Finally, we want to remark that the computational complexity of the estimation boils down to the operations in (16) which are tractable even for larger systems.

C. PARAMETRIC UNCERTAINTY

Theorem 1 allows to efficiently estimate the unknown parameters and the full covariance matrix of the probabilistic multi-step model in (5). While we have the favorable property that the estimated parameters are unbiased, we can, in practice, only consider a finite number of samples for the identification task. Consequently, we must consider the parametric uncertainty of the estimated parameters for which we introduce the following theorem.

Theorem 2: Assumptions 1-4 hold, and we have a non-informative prior distribution for the parameters $\hat{\boldsymbol{W}}$*. We have m samples of data from (14) and maximum likelihood estimates of the parameter and covariance matrix from (16). Considering the posterior distribution of the parameters, i.e. the parametric uncertainty, we have the distributed multi-step prediction:*

$$p(\hat{\boldsymbol{y}}^*_{[1,N]}|\mathcal{D}, \boldsymbol{y}_0, \boldsymbol{u}_{[0,N-1]}) = \mathcal{N}(\hat{\bar{\boldsymbol{y}}}^*_{[1,N]},\ \alpha(\boldsymbol{v})\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}), \tag{26a}$$

with:

$$\hat{\bar{\boldsymbol{y}}}^*_{[1,N]} = \hat{\mathcal{O}}^*_{\boldsymbol{A}}\boldsymbol{y}_0 + \hat{\mathcal{T}}^*_{\boldsymbol{A},\boldsymbol{B}}\boldsymbol{u}_{[0,N-1]}, \tag{26b}$$

$$\alpha(\boldsymbol{v}) = (1 + \boldsymbol{v}^\top\hat{\boldsymbol{\Sigma}}^*_p\boldsymbol{v}), \tag{26c}$$

where $\boldsymbol{v}^\top = [\boldsymbol{y}_0^\top, \boldsymbol{u}_{[0,N-1]}^\top]$.

Proof: We consider the multi-step model in the form of (12) and have noise disturbed independent variables from (13). We have estimated parameters $\hat{\boldsymbol{W}}^*$ and $\hat{\boldsymbol{\Sigma}}_t^*$ from Theorem 1. The posterior of the identified parameters is

$$p(\hat{\boldsymbol{W}}^*|\mathcal{D}, \hat{\boldsymbol{\Sigma}}_t^*) = \frac{p(\mathcal{D}|\hat{\boldsymbol{W}}^*, \hat{\boldsymbol{\Sigma}}_t^*)p(\hat{\boldsymbol{W}}^*)}{p(\mathcal{D}|\hat{\boldsymbol{\Sigma}}_t^*)}. \tag{27}$$

For the likelihood, we again consider the vectorized formulation in (24). With the non-informative prior, i.e. $p(\hat{\boldsymbol{W}}) \sim 1$, we obtain the posterior distribution for the vectorized parameters $\text{vec}(\hat{\boldsymbol{W}}^*)$:

$$p(\text{vec}(\hat{\boldsymbol{W}})|\mathcal{D}, \hat{\boldsymbol{\Sigma}}_t^*) = \mathcal{N}(\text{vec}(\hat{\boldsymbol{W}}), \hat{\boldsymbol{\Sigma}}_t^* \otimes \hat{\boldsymbol{\Sigma}}_p^*). \tag{28}$$

We then vectorize the linear model (10) for a single test point, yielding:

$$\boldsymbol{t}^\top = \boldsymbol{v}^\top\hat{\boldsymbol{W}} + \boldsymbol{e}_t^\top$$
$$\overset{\text{vec}(\cdot)}{\Leftrightarrow} \boldsymbol{t} = \text{vec}\left(\boldsymbol{v}^\top\hat{\boldsymbol{W}}\right) + \boldsymbol{e}_t = \left(\boldsymbol{I} \otimes \boldsymbol{v}^\top\right)\text{vec}\left(\hat{\boldsymbol{W}}\right) + \boldsymbol{e}_t.$$

Considering the distribution of the vectorized parameters in (28), we obtain a multivariate normal distribution for the response variables with mean $\mathbb{E}[\boldsymbol{t}] = \hat{\boldsymbol{W}}^\top \boldsymbol{v}$ and covariance

$$\begin{aligned}\mathrm{Cov}[\boldsymbol{t}] &= (\boldsymbol{I} \otimes \boldsymbol{v}^\top)(\hat{\boldsymbol{\Sigma}}_t^* \otimes \hat{\boldsymbol{\Sigma}}_p^*)(\boldsymbol{I} \otimes \boldsymbol{v}^\top)^\top + \hat{\boldsymbol{\Sigma}}_t^* \\ &= \hat{\boldsymbol{\Sigma}}_t^* \otimes \boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}_p^* \boldsymbol{v} + \hat{\boldsymbol{\Sigma}}_t^*.\end{aligned}$$

The first equality stems from the properties of a linear transformation of a normally distributed random variable [18, 20.23 b], and the second equality follows from the mixed product property of the Kronecker product [18, 11.11]. For a single test point $\boldsymbol{v}$, the term $\boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}_p^* \boldsymbol{v}$ is a scalar and the Kronecker product simplifies to a scalar multiplication:

$$\mathrm{Cov}[\boldsymbol{t}] = (\boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}_p^* \boldsymbol{v})\boldsymbol{\Sigma}_t^* + \hat{\boldsymbol{\Sigma}}_t^* = (\boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}_p^* \boldsymbol{v} + 1)\hat{\boldsymbol{\Sigma}}_t^* = \alpha(\boldsymbol{v})\hat{\boldsymbol{\Sigma}}_t^*.$$

With mean and covariance, we thus have the distribution:

$$p(\boldsymbol{t}|\mathcal{D}, \boldsymbol{v}) = \mathcal{N}(\hat{\boldsymbol{W}}^{*\top} \boldsymbol{v}, \alpha(\boldsymbol{v})\hat{\boldsymbol{\Sigma}}_t^*). \tag{29}$$

From the definition of $\boldsymbol{v}, \boldsymbol{t}$ and considering (16), we obtain the distributed multi-step prediction in (26). □

The distributed multi-step prediction including the parametric uncertainty is thus given in (26). We observe the following behavior as the number of samples m used for the identification increases. As shown in [21, 3.3.2], we have for $m \rightarrow \infty$ that $\alpha(\boldsymbol{v}) \rightarrow 1$. In the same limit, the identified parameters and covariance matrix converge to their expectation in (17) and we recover exactly the distribution (5). We demonstrate this behavior in the numerical example in Section VI.

IV. STOCHASTIC MPC WITH IDENTIFIED MULTI-STEP MODEL

With the distribution in (26) we have the data-based equivalent to (5) including the parametric uncertainty from the identification task. The deterministic formulation of the stochastic optimal control problem (9) with identified multi-step model is given as:

$$\begin{aligned}\min_{\boldsymbol{u}_{[0,N-1]} \in \mathbb{A}} \quad & \|\hat{\boldsymbol{y}}^*_{[1,N]}\|^2_{\boldsymbol{Q}} + \|\boldsymbol{u}_{[0,N-1]}\|^2_{\boldsymbol{R}} \\ & + \mathrm{trace}\left(\alpha(\boldsymbol{v})\boldsymbol{Q}\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}\right) \qquad \text{(30a)}\\ \text{s.t.} \quad & \boldsymbol{a}_j^\top \hat{\boldsymbol{y}}^*_{[1,N]} \le b_j - c_p(\epsilon)\|\boldsymbol{a}_j\|_{\alpha(\boldsymbol{v})\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}} \quad \forall j \in \mathbb{I}_{[1,n_c]}. \qquad \text{(30b)}\end{aligned}$$

The weight matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, constraint set $\mathbb{A}$ and the chance constraints, defined with $\boldsymbol{a}_j$, b_j $\forall j \in \mathbb{I}_{[1,n_c]}$ and ϵ, are analogous to (9). As a main difference between (9) and (30), the mean $\hat{\boldsymbol{y}}^*_{[1,N]}$ and covariance $\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}$ now stem from distribution (26). The parametric uncertainty of the identified model results in the factor $\alpha(\boldsymbol{v})$, introduced in (26c), where $\boldsymbol{v}^\top = [\boldsymbol{y}_0^\top, \boldsymbol{u}_{[0,N-1]}^\top]$.

Due to the parametric uncertainty, and in contrast to (9), the optimization problem in (30) is not a quadratic problem anymore. We show in Appendix A that problem (30) constitutes a convex second-order cone program [19]. In contrast to previous work [8], where the parametric uncertainty is included as an ellipsoidal uncertainty set, we propose to solve problem (30) directly.

A. RECURSIVE APPLICATION

The stochastic optimal control problem in (30) is solved repeatedly at each sampling time to yield closed-loop control actions. This introduces the well known challenges to guarantee recursive feasibility and stability [14]. Additionally, we have to consider a limitation of the identified multi-step model. As shown in Theorem 1, the multi-step model identifies the parameters and covariance matrix for the distribution in (5). This distribution incorporates a Kalman filter correction step, considering the prior distribution of the initial state $p_{\text{prior}}(\boldsymbol{x}_0)$ and the current measurement. However, this prior for the initial state distribution is not explicitly formulated. Instead, it coincides with the distribution of the initial state from the sampled data $p_{\text{samp.}}(\boldsymbol{x}_0)$, as required for Theorem 1. In other words, the multi-step model always considers the same prior distribution, that is, $p_{\text{prior}}(\boldsymbol{x}_0) = p_{\text{samp.}}(\boldsymbol{x}_0)$.

This has two important limitations. First, it is imperative that the multi-step model is evaluated only for initial states $\boldsymbol{x}_{0,\text{eval}}$ that are reasonably probable in terms of the distribution of the initial state from the sampled data. This is an intuitive limitation, as it is detrimental to perform a Kalman filter correction step with poorly selected prior state distribution. It is, however, a limitation in the sense that data for the identification task must be sampled in the same range that is expected for the closed-loop operation.

The second limitation arises from the fact that with identified multi-step model the prior state distribution is not updated. Updating the prior distribution is typically an important step of Kalman filtering and allows for convergence to the true state distribution over multiple iterations. In practice, this limitation increases the uncertainty of the predicted future measurements and may lead to more conservative control actions of the SMPC controller.

V. STATE-SPACE IDENTIFICATION

To have a comparative baseline for our proposed method, we also employ a probabilistic state-space identification approach. We consider the same data for the identification task which stems from system (1) with unknown parameters and process and measurement noise covariances. In particular, we have the data with individual samples:

$$\boldsymbol{v}_i = \left[\boldsymbol{y}_0^{(i)\top}, \ \boldsymbol{u}_0^{(i)\top}\right]^\top, \ \boldsymbol{t}_i = \boldsymbol{y}_1^{(i)} \quad \forall i \in \mathbb{I}_{[1,m]}. \tag{31}$$

A. CHALLENGES WITH STATE-SPACE IDENTIFICATION

We can identify a state-space model as a special case of the described multi-step identification approach with $N = 1$. In that case, it follows from (3) that $\mathcal{O}(\boldsymbol{A}) = \boldsymbol{A}$ and $\mathcal{T}(\boldsymbol{A}, \boldsymbol{B}) = \boldsymbol{B}$. Unfortunately, state-space identification poses two major challenges which both stem from the requirement that the identified model must be recursively evaluated to obtain the multi-step prediction.

The first challenge is related to the measurement noise of the system. From Theorem 1, we obtain parameters $\hat{\boldsymbol{A}}^*$, $\hat{\boldsymbol{B}}^*$ and $\hat{\boldsymbol{\Sigma}}^*_{y,1}$ with the properties:

$$\mathbb{E}[\hat{\boldsymbol{A}}^*] = \boldsymbol{A}\boldsymbol{L},$$
$$\mathbb{E}[\hat{\boldsymbol{B}}^*] = \boldsymbol{B},$$
$$\mathbb{E}[\hat{\boldsymbol{\Sigma}}^*_{y,1}] = \boldsymbol{\Sigma}_{y,1} + \boldsymbol{A}(\boldsymbol{\Sigma}_{x,0} + \boldsymbol{L}\boldsymbol{\Sigma}_{x,0})\boldsymbol{A}^\top,$$

where

$$\boldsymbol{\Sigma}_{y,1} = \boldsymbol{A}\boldsymbol{E}_x\boldsymbol{\Sigma}_x\boldsymbol{E}_x^\top\boldsymbol{A}^\top + \boldsymbol{E}_y\boldsymbol{\Sigma}_y\boldsymbol{E}_y^\top. \tag{32}$$

In expectation, the identified parameters are thus identical to the parameters in (6). With:

$$\hat{\bar{\boldsymbol{y}}}_1^* = \hat{\boldsymbol{A}}^*\boldsymbol{y}_0 + \hat{\boldsymbol{B}}^*\boldsymbol{u}_0, \tag{33}$$

we thus have the conditional distribution (5) for $N = 1$:

$$p(\hat{\boldsymbol{y}}_1^*|\boldsymbol{y}_0, \boldsymbol{u}_0) = \mathcal{N}(\hat{\bar{\boldsymbol{y}}}_1^*, \hat{\boldsymbol{\Sigma}}^*_{y,1}). \tag{34}$$

We can recursively evaluate (34) to obtain a distributed multi-step prediction $\hat{\boldsymbol{y}}_{[1,N]}$. Unfortunately, this will not yield the correct multi-step prediction in (5) for $N > 1$. This is due to the fact that (32) contains contributions from both the process noise and the measurement noise. The recursive evaluation of (34) thus propagates the measurement noise which can significantly inflate the uncertainty of the multi-step prediction.

The second challenge with state-space identification is the parametric uncertainty discussed in Subsection III-C. This uncertainty can be considered for a single-step prediction with (26) and $N = 1$. However, there is no closed-form solution for the recursive application of (34) if parametric uncertainties are considered.

B. IDENTIFIED STATE-SPACE MODEL FOR SMPC

The distributed multi-step system response obtained from the recursively evaluated state-space model is expected to have significant shortcomings due to the challenges discussed above. Regardless, we suggest an approach to formulate the stochastic MPC problem with the identified state-space model. The approach serves as an important comparative baseline, as it is obtained under the same premises as for the multi-step identification.

We formulate the stochastic MPC problem as in (9), using the identified state-space model with parameters $\hat{\boldsymbol{A}}^*$, $\hat{\boldsymbol{B}}^*$ and covariance $\hat{\boldsymbol{\Sigma}}^*_{y,1}$. The distributed multi-step prediction is obtained by recursively evaluating (34) without considering the parametric uncertainty.

VI. LINEAR CASE STUDY

In this and the following section, we investigate the proposed method in two simulation studies. We first consider a linear system with noisy state-feedback and investigate a nonlinear system with noisy output-feedback in Section VII. For both simulation studies, the complete code and our results are available online.[1]

A. SYSTEM DESCRIPTION

We consider the linear building model previously presented in [24]. The system has five states $\boldsymbol{x} = [T_1, \ldots, T_4, T_a]^\top$, where T_i [°C] is the temperature in room i and T_a [°C] the ambient temperature. The room temperatures are controlled with combined heating and cooling units $\boldsymbol{u} = [\dot{Q}_1, \ldots, \dot{Q}_4]^\top$ [kW]. The dynamics are modeled as a resistance network [25]. We assume that an uncertain forecast T_f [°C] of the ambient temperature is available, and model this situation with $\dot{T}_a = \tau_a(T_f - T_a) + e_{T,a}$, and significant process noise. The model is parameterized as in [24], including $\tau_a = 1/72.000$ and discretized with timestep $\Delta t = 3600$ s. Furthermore, we define the variances $\sigma_x^\top = [0, 0, 0, 0, 0.5]$ and $\sigma_y^\top = 10^{-1} \cdot [1, 1, 1, 1, 1]$ for the process and measurement noise. The covariance matrices are then obtained as $\boldsymbol{\Sigma}_x = \mathrm{diag}(\sigma_x^2)$ and $\boldsymbol{\Sigma}_y = \mathrm{diag}(\sigma_y^2)$.

B. SYSTEM IDENTIFICATION

In our first investigation, we seek to show that with the proposed multi-step identification, we obtain a data-based model that represents the distributed system response in (5). We want to highlight, in particular, that this distribution incorporates the effect of the state estimation as discussed in Subsection II-B.

For the investigation, we first determine an initial state distribution according to Assumption 3. With standard deviation $\sigma_{x,0}^\top = [2, 2, 2, 2, 5]$ and mean $\bar{\boldsymbol{x}}_0^\top = [20, 20, 20, 20, 15]$, we have:

$$\boldsymbol{x}_0 \sim \mathcal{N}(\bar{\boldsymbol{x}}_0, \mathrm{diag}(\sigma_{x,0}^2)). \tag{35}$$

Two multi-step models, both with horizon $N = 12$, are identified. We denote MSM_a for the identified multi-step model using $m = 1000$ sequences, and MSM_b for the identified multi-step model with $m = 100$ sequences. All sequences start from a random initial state according to distribution (35) and we normalize the data to achieve a zero-mean as required for Assumption 3.

Additionally, we obtain the ground truth for the predictive distribution. To this end, we consider the true system and covariance matrices and evaluate distribution (5), which includes the Kalman filter correction step for the prior distribution in (35).

In Figure 1, we display the distributed multi-step prediction $p(\hat{\boldsymbol{y}}_{[1,N]})$ over the prediction horizon. We show the mean and the standard deviation for T_1, T_2 and T_a. The other states are omitted from the plot to improve readability. Furthermore, we show the joint distribution $p(T_i, T_j)$ at $t = 12$ h. In comparison to the ground truth, MSM_a yields almost the identical predictive distribution. This is consistent with our theoretical results presented in Theorem 1. Due to the large number of samples the identified parameters are almost

[1] https://github.com/4flixt/2023_Stochastic_MSM

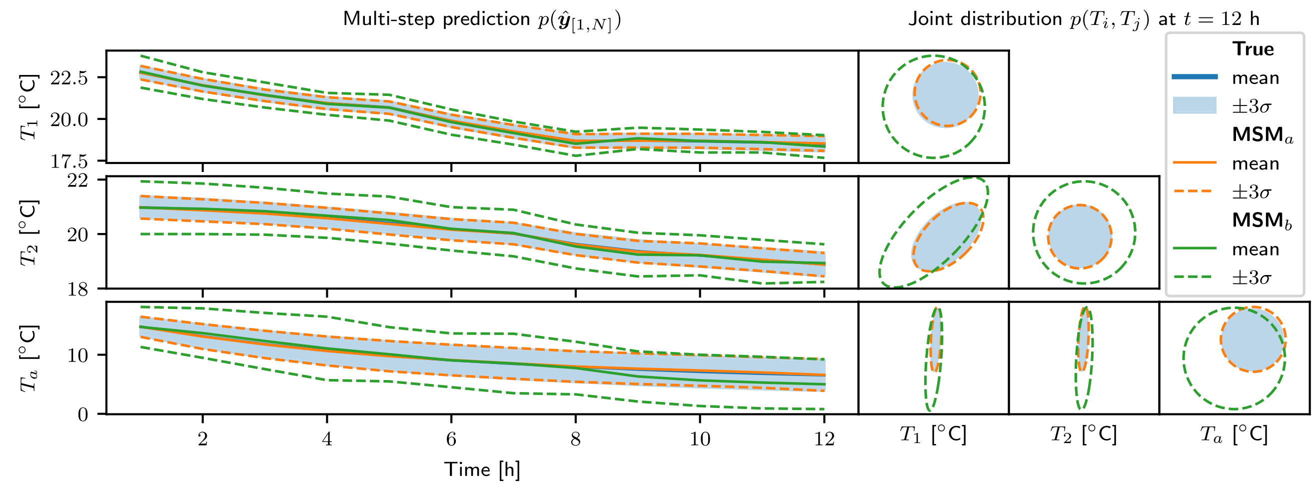


FIGURE 1. Distributed multi-step prediction of the building system for an initial measurement y_0 and random control sequence $u_{[0,N-1]}$ for $N = 12$. Comparison of true distribution (5) and two identified multi-step models (MSM). The distribution of MSM$_a$ with $m = 1000$ samples and MSM$_b$ with $m = 100$ samples are obtained according to (26).

identical to their expectation in (17), and we recover the true distribution in (5) from data. On the other hand, MSM$_b$ is identified with only $m = 100$ samples and therefore experiences significant parametric uncertainty, according to Theorem 2. However, while the obtained distribution is broader and the mean is shifted, it still approximates the true distribution in (5).

C. STOCHASTIC MPC

We formulate a stochastic MPC controller based on the identified multi-step model (MSM-SMPC). As a comparative baseline, we also formulate a stochastic MPC controller based on an identified state-space model (SSM-SMPC). Both models are obtained from the same data as described in the previous subsection.

For the control objective, we choose the cost function:

$$J(\boldsymbol{u}_{[0,N-1]}) = \sum_{k=0}^{N-2} (5\|\boldsymbol{u}_k\|_2^2 + 10\|\boldsymbol{u}_{k+1} - \boldsymbol{u}_k\|_2^2) + \|\boldsymbol{u}_{N-1}\|_2^2.$$

The actuators (heating and cooling power) are limited to $-6\,\text{kW} \leq \dot{Q}_i \leq 6\,\text{kW}$ for all rooms. We have the individual chance constraints $P(T_i \geq 18) \geq (1 - \epsilon)$, which can be violated with probability $\epsilon = 10^{-3}$. The stochastic optimal control problem is implemented and solved using CasADi [26] with IPOPT [27].

In a first investigation, we compare the open-loop predictions of both controllers. We define a scenario with initial state $\boldsymbol{x}_0 = [23, 20, 20, 20, 10]^\top$°C and corresponding noise disturbed measurement $\boldsymbol{y}_0$. We then solve the optimal control problem in (30) for the multi-step model and (9) for the state-space model. As discussed in Section V, the SMPC controller based on the state-space model does not consider the parametric uncertainty.

TABLE 1. Building system: Closed-loop performance over a period of 50h with mean and standard deviation computed over 20 samples.

	MSM-SMPC	SSM-SMPC
$\sum_i Q_i$ [kWh]	640.9±32.4	698.8±29.9
cons. viol. [%]	0.0±0.0	0.0±0.0

The resulting open-loop predictions are shown in Figure 2. We display the predicted mean $\hat{\boldsymbol{y}}_{[1,N]}$ and standard deviation, considering $c_p(\epsilon)$ for the chosen violation probability. Furthermore, we show 50 samples of the true system response for the optimal open-loop input sequence, considering the same initial state and randomly drawn process and measurement noise. It can be seen in Figure 2 that the MSM-SMPC controller yields a suitable sequence of inputs to minimize the cost function while satisfying the chance constraints. While the chance constraints are also satisfied for the SSM-SMPC controller, the predicted uncertainty bounds are much larger and overly conservative. In the investigated scenario, with forecasted ambient temperature of $T_f = 10$°C, this leads to suboptimal controls action with unnecessarily high energy consumption. This effect is expected, as discussed in Section V, as the identified state-space model incorrectly propagates the measurement noise.

In a second investigation, we compare the closed-loop performance of the two controllers. To this end, we consider the same initial state as in the previous section and recursively control the system for a period of 50 h with either the MSM-SMPC or SSM-SMPC controller. The samples differ in the randomly drawn sequences of process and measurement noise. Finally, we obtain the percentage of constraint violations and the total energy consumption for each run and compute mean and standard deviation of these values. The performance metrics are shown in Table 1. The MSM-SMPC controller yields on average 8.3 % less energy consumption

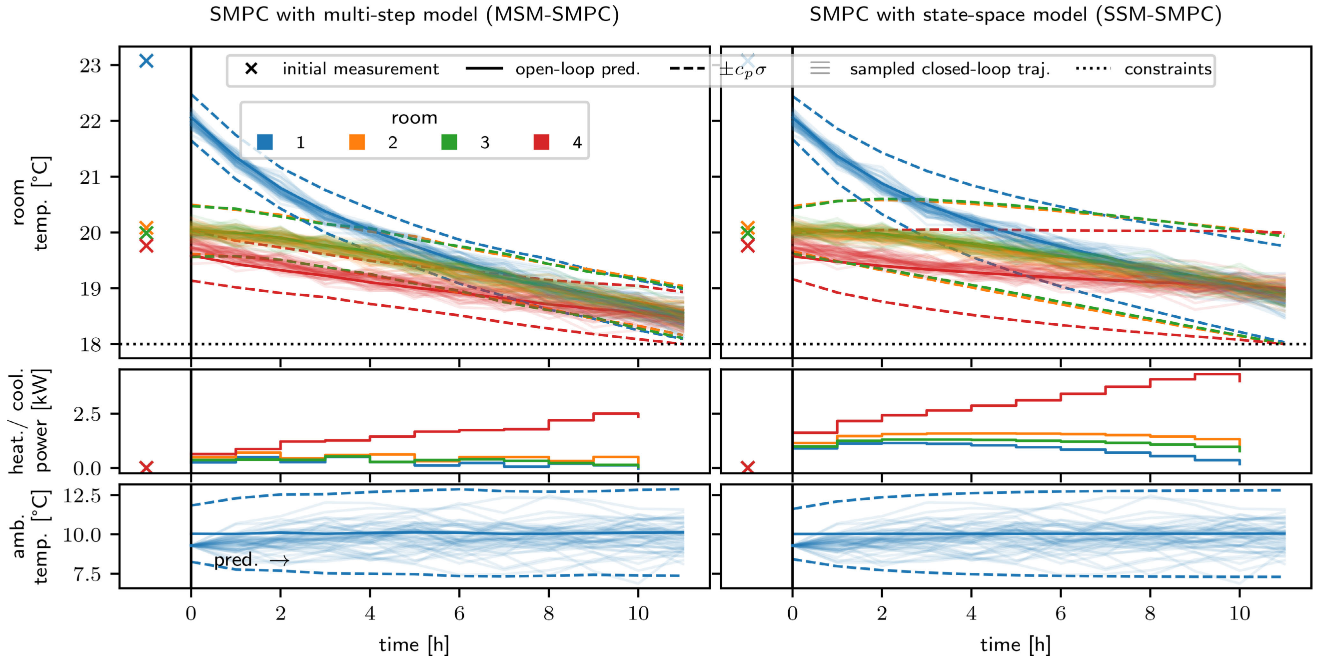


FIGURE 2. Building system: Comparison of open-loop prediction (mean and predicted standard deviation) of SMPC with multi-step model and SMPC with state-space model for the same initial measurement. The standard deviation is scaled with c_p introduced in (9). Using the open-loop control inputs, 50 samples of the true system response are drawn.

than the SSM-SMPC controller. Both controllers satisfy the chance constraints with no violations.

VII. NONLINEAR CASE STUDY WITH OUTPUT-FEEDBACK

The proposed multi-step identification approach is derived for linear systems with noisy state-feedback. Additionally, we explore its applicability to a nonlinear system with output-feedback in this section. The application to nonlinear systems is motivated by interpreting the process noise in (1a) as an additive nonlinear term. While this interpretation violates Assumption 2, that is, the process noise is not normally distributed, we expect to identify a multi-step model where the identified uncertainty approximately encompasses the nonlinearities.

As in the previous section, we compare the SMPC based on an identified multi-step model with a variant based on state-space identification.

A. SYSTEM DESCRIPTION

For the nonlinear case study, we consider the continuously stirred tank reactor (CSTR) previously introduced in [28]. The reactor is modeled with four states $x^\top = [c_A, c_B, T_R, T_k]$, with the concentrations c_A [$\mathrm{mol\,L^{-1}}$] and c_B [$\mathrm{mol\,L^{-1}}$] and the temperatures T_R [°C] of the reactor and T_K [°C] of the cooling jacket. As control inputs, we have the flow rate $\dot{V}$ [$\mathrm{m^3\,s^{-1}}$] which is normalized with the reactor volume V_R [$\mathrm{m^3}$], yielding $F = \dot{V}/V_R$, and the heat removed from the jacket $\dot{Q}$ [$\mathrm{kJ\,h^{-1}}$]. For the safe operation of the CSTR, the states and inputs must lie within the bounds shown in Table 2. In our investigated scenario, the nonlinear

TABLE 2. Bounds for the CSTR system. [1] Chance constraint for SMPC.

	states				inputs	
	c_A $\mathrm{mol\,L^{-1}}$	c_B $\mathrm{mol\,L^{-1}}$	T_R °C	T_K °C	F $\mathrm{h^{-1}}$	$\dot{Q}$ $\mathrm{kJ\,h^{-1}}$
lower	0.1	0.1	50.0	50.0	5.0	-8500.0
upper	2.0	2.0	135.0[1]	140.0	60.0	0.0

CSTR system experiences measurement noise with standard deviation $\sigma_y^\top = [0.01, 0.01, 0.5, 0.5]$ but no additional process noise.

The nonlinear model is created and simulated in do-mpc [29] with a timestep of 18 s. The model equations and parameters can be found in [28] or online.[2]

B. SYSTEM IDENTIFICATION

As in the previous section, we compare the identified state-space model (SSM) and multi-step model (MSM). Furthermore, we also investigate the effect of having full state-feedback vs. output-feedback for system identification and the successive control application. As in related data-based approaches [5], [6], [9], [11], output-feedback can be incorporated in the proposed methods by introducing the state:

$$x_k^\top = [y_{[k-l,k]}^\top, u_{[k-l,k-1]}^\top]. \tag{36}$$

If the system is observable, and with l chosen larger than the system lag, the introduction of state (36) allows to

[2]https://github.com/4flixt/2023_Stochastic_MSM

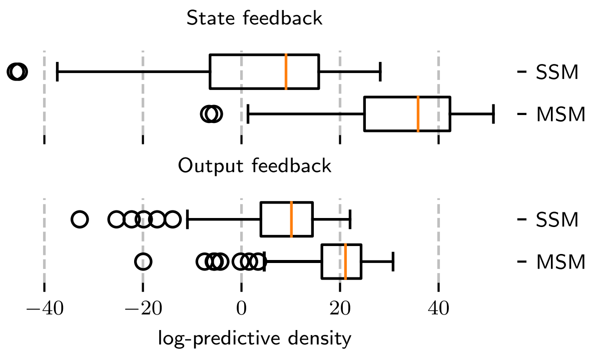


FIGURE 3. Nonlinear CSTR system: Comparison of log-predictive density for the identified state-space model (SSM) vs. the multi-step model (MSM). The predictive density is computed for 100 test cases as shown in (37). Representation as box-plot with filtered outliers, inter-quartile range, minimum-maximum range and median.

reformulate the system with output-feedback in the form of (1) [11] and satisfies Assumption 1. Unfortunately, considering output-feedback, by introducing the state in (36), violates Assumption 2 as the additive measurement noise on the newly introduced state (36) is now correlated. While this correlation is known to yield biased estimates [3, Sec. 5], we show in the following example that the proposed method can still lead to good results.

For the system identification, we gather $m = 500$ simulated sequences of length $L = N + l$ with $N = 20$ and $l = 1$ for the models with state-feedback and $l = 3$ for model with output-feedback. For both cases we also create $m = 100$ sequences of test data. All sampled sequences are created with uniformly random initial state, within the bounds shown in Table 2, and with persistently exciting random inputs.

In contrast to the linear case study in Section VI, there is no ground-truth distribution to evaluate the performance of the identified models. We therefore investigate the quality of the obtained probabilistic models by computing the logarithm of the predicted distribution, that is:

$$\log p\left(\boldsymbol{y}_{[1,N]} = \boldsymbol{y}^{(i)}_{[1,N]} | \mathcal{D}, \boldsymbol{y}^{(i)}_0, \boldsymbol{u}^{(i)}_{[0,N-1]}\right), \tag{37}$$

for the test samples $\left(\boldsymbol{y}^{(i)}_{[0,N]}, \boldsymbol{u}^{(i)}_{[0,N-1]}\right)$ $\forall i \in \mathbb{I}_{[1,m]}$. The log-predictive density is an expressive measure for the quality of the identified probabilistic model. High values indicate high confidence in correct predictions, while low values indicate high confidence in incorrect predictions.

We display the log-predictive density in Figure 3 in the form of a box-plot to visualize the median, the quartile range, minimum and maximum as well as outliers. The identified multi-step model shows clear advantages over the recursively evaluated state-space model. Both, for state-feedback and output-feedback, the median and quartile ranges of the log-predictive density are significantly higher for the multi-step model. We also see that the log-predictive density obtained with the MSM is increased in the case of state-feedback. The state-space model, on the other hand, experiences more severe outliers for state-feedback, reducing the overall log-predictive density. We reason that with state-feedback the uncertainty in the predictive distribution is reduced as more information is available. However, this only benefits the log-predictive density of the MSM which yields accurate predictions with high confidence in that case.

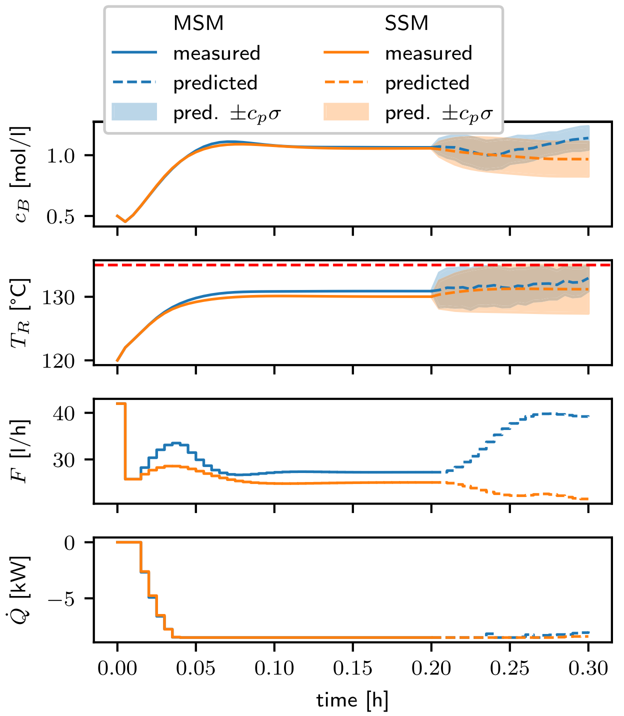


FIGURE 4. Nonlinear CSTR: Output-feedback SMPC closed-loop results and future prediction with uncertainty bounds. Comparison of SMPC and prediction with identified multi-step model (MSM) and identified state-space model (SSM) for an exemplary initial condition.

TABLE 3. Nonlinear CSTR system: Closed-loop SMPC performance over 50 experiments. Comparison of MSM and SSM models with state or output-feedback.

Output-feedback	MSM	product [mol]	6.15±0.23
		max. cons. viol [°C]	0.00±0.00
	SSM	product [mol]	5.79±1.10
		max. cons. viol [°C]	0.01±0.06
State-feedback	MSM	product [mol]	7.17±0.21
		max. cons. viol [°C]	0.00±0.00
	SSM	product [mol]	6.16±0.21
		max. cons. viol [°C]	0.00±0.00

C. STOCHASTIC MPC

We proceed our investigation with an analysis of the stochastic MPC controllers obtained with the multi-step model (MSM-SMPC) and the state-space model (SSM-SMPC). Furthermore, we continue to investigate the differences between output-feedback and state-feedback for the identified systems and the control application.

All investigated controller variants are implemented with the control objective to maximize the production of component B. Furthermore, we seek to penalize rapid changes of the control inputs. To this end, we propose the cost function:

$$J(\boldsymbol{y}_{[1,N]}, \boldsymbol{u}_{[0,N-1]}) = -\sum_{k=1}^{N} c_{B,k} + \sum_{k=1}^{N-1} \left(10^{-2}\Delta F_K + 10^{-5}\Delta \dot{Q}_k\right),$$

with $\Delta F_k = F_k - F_{k-1}$ and $\Delta \dot{Q}_k = \dot{Q}_k - \dot{Q}_{k-1}$. The problem is formulated with horizon $N = 20$ and must consider the constraints in Table 2. Of those constraints, we formulate the safety critical bound of the reactor temperature as a chance constraint $P(T_R \leq 135) \geq (1-\epsilon)$, with probability of violation $\epsilon = 10^{-3}$.

To illustrate the behavior of the MSM-SMPC and SSM-SMPC controllers, we showcase an exemplary closed-loop trajectory in Figure 4, for the case of output-feedback. Both controllers run for 40 timesteps, corresponding to 12 min of simulation time. Additionally, we display mean and standard deviation, scaled with $c_p(\epsilon)$, of the open-loop prediction for the final timestep. Both controllers achieve satisfactory control performance in the example, yielding a high concentration of c_B while safely avoiding the constraint $T_R \leq 135$. However, we see in Figure 4 that the SMPC controller with multi-step model can operate significantly closer to the chance constraint for the reactor temperature. This allows to realize a higher product concentration c_B and higher normalized flow rate F, which yields overall more product.

To further quantify the performance of SSM-MPC and MSM-SMPC, also for the case of state-feedback, we present the results in Table 3. We display the amount of produced component B and the maximum constraint violation of the chance constraint $T_R \leq 135$ for all investigated controller variants. These performance indicators are computed as mean and standard deviation for 50 independent experiments with different initial state sampled uniformly within the bounds in Table 2. The table supports the qualitative finding from Figure 4. MSM-SMPC outperforms SSM-SMPC in the scenario with state-feedback and for the scenario with output-feedback. For the output-feedback scenario, we also see that SSM-SMPC leads to minor constraint violations. Finally, we observe that both controllers have a better performance, in terms of the obtained product, when state-feedback is available. In the case of state-feedback the MSM-SMPC controllers results, on average, in 16.4 % more product than the SSM-SMPC controller.

VIII. CONCLUSION

In this work, we propose a novel approach that combines probabilistic multi-step system identification with stochastic Model Predictive Control (SMPC). Our identification procedure is derived for linear systems with noisy state-feedback and with Gaussian process and measurement noise. In contrast to previous work, our proposed method does not require knowledge of the noise covariance matrices.

As a main contribution, we derive that the identified multi-step model yields, in expectation, the true distribution of the future measurements. We show that evaluating our identified model with noisy state-measurements is equivalent to estimating the initial state distribution and propagating this distribution with the known system dynamics. In this way, the identified multi-step model performs an implicit state estimation and can directly be used to formulate an SMPC problem for noisy state-measurements.

We demonstrate the theoretical findings and the performance of our proposed data-based SMPC controller in two simulation studies. In comparison to a SMPC controller based on an identified state-space model, we achieve significantly better performance and safer operation. Furthermore, we showcase in the second simulation study that our proposed method can also be applied to a nonlinear system with output-feedback, despite not being originally derived for this context.

In future work, we seek to rigorously extend our method to linear systems with output-feedback, by considering the correlation of the measurement noise. Furthermore, we seek to extend the method by updating the identified multi-step model with new data in a recursive fashion.

APPENDIX

A. CONVEXITY OF SMPC PROBLEM WITH PARAMETRIC UNCERTAINTY

We show that problem (30) is a convex optimization problem. The last term in the objective function in (30a) can be reformulated as:

$$\text{trace}\left(\alpha(\boldsymbol{v})\boldsymbol{Q}\boldsymbol{\Sigma}^*_{y,[1,N]}\right) = \left(1 + \boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}^*_p \boldsymbol{v}\right) \text{trace}\left(\boldsymbol{Q}\boldsymbol{\Sigma}^*_{y,[1,N]}\right),$$

with positive definite matrix $\hat{\boldsymbol{\Sigma}}^*_p$ due to Assumption 4. Adding the convex term to the overall objective function does not change the convexity of the problem. The expression $\hat{\boldsymbol{y}}^*_{[1,N]}$ in (26b) is linear in the optimization variables $\boldsymbol{u}_{[0,N-1]}$. It remains to show that the inequality constraint (30b) is convex. To this end, we show in the following lemma that (30b) represents a second order cone and is thus convex [19].

Lemma 2: Let Assumption 4 hold. The constraint (30b) has the form $\forall j \in \mathbb{I}_{[1,n_c]}$:

$$\boldsymbol{g}_j(\boldsymbol{v}) \leq 0,$$

with:

$$\boldsymbol{g}_j(\boldsymbol{v}) = \tilde{\boldsymbol{a}}_j^\top \boldsymbol{v} + \sqrt{\boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}^*_p \boldsymbol{v} + 1} - \tilde{b}_j, \tag{38}$$

where:

$$\tilde{\boldsymbol{a}}_j^\top = \boldsymbol{a}_j^\top [\hat{\mathcal{T}}^*_{\boldsymbol{A},\boldsymbol{B}},\ \hat{\mathcal{O}}^*_{\boldsymbol{A}}] c_p(\epsilon)^{-1} d_j^{-1/2}, \tag{39a}$$

$$\tilde{b}_j = b_j c_p(\epsilon)^{-1} d_j^{-1/2}, \tag{39b}$$

and:

$$d_j = \boldsymbol{a}_j^\top \hat{\boldsymbol{\Sigma}}^*_{y,[1,N]} \boldsymbol{a}_j. \tag{40}$$

The constraint represents a second order cone.

Proof: To obtain the form in (38), we first insert (26b) in (30b) and consider the definition of $\boldsymbol{v}$:

$$\boldsymbol{a}_j^\top [\hat{\mathcal{T}}^*_{\boldsymbol{A},\boldsymbol{B}}, \ \hat{\mathcal{O}}^*_{\boldsymbol{A}}]\boldsymbol{v} \leq b_j - c_p(\epsilon)\|\boldsymbol{a}_j\|_{\alpha\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}}. \tag{41}$$

We expand the term $\|\boldsymbol{a}_j\|_{\alpha\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}}$ and obtain:

$$\begin{aligned} \|\boldsymbol{a}_j\|_{\alpha\hat{\boldsymbol{\Sigma}}^*_{y,[1,N]}} &= \sqrt{\alpha \boldsymbol{a}_j^\top \hat{\boldsymbol{\Sigma}}^*_{y,[1,N]} \boldsymbol{a}_j} = \sqrt{d_j}\sqrt{\alpha}, \\ &= \sqrt{d_j}\sqrt{1 + \boldsymbol{v}^\top \hat{\boldsymbol{\Sigma}}^*_p \boldsymbol{v}}, \end{aligned} \tag{42}$$

with constant d_j introduced in (40). We insert (42) in (41) and divide by $\sqrt{d_j} c_p(\epsilon)$. By introducing and substituting the expressions (39), we obtain the desired expression (38). □

REFERENCES

[1] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, *Model Predictive Control: Theory, Computation, and Design*, 2nd ed. San Francisco, CA, USA: Nob Hill, 2017.

[2] M. Morari and J. H. Lee, "Model predictive control: Past, present and future," *Comput. Chem. Eng.*, vol. 23, nos. 4–5, pp. 667–682, May 1999.

[3] K. J. Åström and P. Eykhoff, "System identification—A survey," *Automatica*, vol. 7, no. 2, pp. 123–162, 1971.

[4] L. Ljung, "System identification," in *Signal Analysis and Prediction* (Applied and Numerical Harmonic Analysis). Basel, Switzerland: Birkhauser, 1998, pp. 163–173.

[5] W. Favoreel, B. D. Moor, and M. Gevers, "SPC: Subspace predictive control," *IFAC Proc. Volumes*, vol. 32, no. 2, pp. 4004–4009, Jul. 1999.

[6] E. Terzi, L. Fagiano, M. Farina, and R. Scattolini, "Learning-based predictive control for linear systems: A unitary approach," *Automatica*, vol. 108, Oct. 2019, Art. no. 108473.

[7] E. Terzi, M. Farina, L. Fagiano, and R. Scattolini, "Robust multi-rate predictive control using multi-step prediction models learned from data," *Automatica*, vol. 136, Feb. 2022, Art. no. 109852.

[8] J. Köhler, K. P. Wabersich, J. Berberich, and M. N. Zeilinger, "State space models vs. multi-step predictors in predictive control: Are state space models complicating safe data-driven designs?" in *Proc. IEEE 61st Conf. Decis. Control (CDC)*, Dec. 2022, pp. 491–498.

[9] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *Proc. 18th Eur. Control Conf. (ECC)*, Jun. 2019, pp. 307–312.

[10] I. Markovsky and F. Dörfler, "Behavioral systems theory in data-driven analysis, signal processing, and control," *Annu. Rev. Control*, vol. 52, pp. 42–64, Jan. 2021.

[11] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "On the design of terminal ingredients for data-driven MPC," *IFAC-PapersOnLine*, vol. 54, no. 6, pp. 257–263, 2021.

[12] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct and indirect data-driven control formulations via regularizations and relaxations," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 883–897, Feb. 2023.

[13] F. Fiedler and S. Lucia, "On the relationship between data-enabled predictive control and subspace predictive control," in *Proc. Eur. Control Conf. (ECC)*, Jun. 2021, pp. 222–229.

[14] A. Mesbah, "Stochastic model predictive control: An overview and perspectives for future research," *IEEE Control Syst. Mag.*, vol. 36, no. 6, pp. 30–44, Dec. 2016.

[15] M. Farina, L. Giulioni, and R. Scattolini, "Stochastic linear model predictive control with chance constraints – a review," *J. Process Control*, vol. 44, pp. 53–67, Aug. 2016.

[16] M. Fochesato and J. Lygeros, "Data-driven distributionally robust bounds for stochastic model predictive control," in *Proc. IEEE 61st Conf. Decis. Control (CDC)*, Dec. 2022, pp. 3611–3616.

[17] L. Hewing, J. Kabzan, and M. N. Zeilinger, "Cautious model predictive control using Gaussian process regression," *IEEE Trans. Control Syst. Technol.*, vol. 28, no. 6, pp. 2736–2743, Nov. 2020.

[18] G. A. Seber, *A Matrix Handbook for Statisticians*. Hoboken, NJ, USA: Wiley, 2008.

[19] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[20] L. Blackmore, M. Ono, and B. C. Williams, "Chance-constrained optimal path planning with obstacles," *IEEE Trans. Robot.*, vol. 27, no. 6, pp. 1080–1094, Dec. 2011.

[21] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.

[22] W. A. Fuller, *Measurement Error Models*. Hoboken, NJ, USA: Wiley, 2009.

[23] J. P. Buonaccorsi, *Measurement Error: Models, Methods, and Applications*. Boca Raton, FL, USA: CRC Press, 2010.

[24] L. Hewing, K. P. Wabersich, and M. N. Zeilinger, "Recursively feasible stochastic model predictive control using indirect feedback," *Automatica*, vol. 119, Sep. 2020, Art. no. 109095.

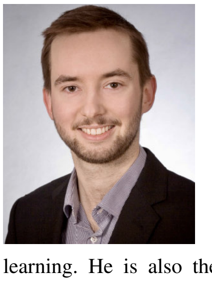
[25] B. Bueno, L. Norford, G. Pigeon, and R. Britter, "A resistance-capacitance network model for the analysis of the interactions between the energy performance of buildings and the urban climate," *Building Environ.*, vol. 54, pp. 116–125, Aug. 2012.

[26] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: A software framework for nonlinear optimization and optimal control," *Math. Program. Comput.*, vol. 11, no. 1, pp. 1–36, Mar. 2019.

[27] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math. Program.*, vol. 106, no. 1, pp. 25–57, Mar. 2006.

[28] K.-U. Klatt and S. Engell, "Gain-scheduling trajectory control of a continuous stirred tank reactor," *Comput. Chem. Eng.*, vol. 22, nos. 4–5, pp. 491–502, Jan. 1998.

[29] F. Fiedler, B. Karg, L. Lüken, D. Brandner, M. Heinlein, F. Brabender, and S. Lucia, "Do-MPC: Towards FAIR nonlinear and robust model predictive control," *Control Eng. Pract.*, vol. 140, Nov. 2023, Art. no. 105676.



FELIX FIEDLER received the M.Sc. degree in chemical engineering from TU Berlin University, Germany, in 2018. He was Ph.D. candidate at TU Berlin University with the HEIBRiDS Graduate School program, until 2020. In the same year, he joined TU Dortmund University, Germany, where he is currently pursuing his Ph.D. with the Chair of Process Automation Systems. His research interest include model predictive control under uncertainty, system identification and machine learning. He is also the current lead developer of the open-source do-mpc toolbox which enables rapid non-linear MPC developments in Python (www.do-mpc.com).

SERGIO LUCIA (Member, IEEE) received the M.Sc. degree in electrical engineering from the University of Zaragoza, Spain, in 2010, and the Dr. (Ing.) degree in optimization and automatic control from TU Dortmund University, Germany, in 2014. He joined Otto-von-Guericke Universität Magdeburg and visited the Massachusetts Institute of Technology as a Postdoctoral Fellow. He was an Assistant Professor with TU Berlin, between 2017 and 2020. Since 2020, he has been a Professor with TU Dortmund University and the Head of the Chair of Process Automation Systems. His research interests include decision-making under uncertainty, distributed control, and the interplay between machine learning techniques and control theory. He is currently an Associate Editor of the *Journal of Process Control*.