2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Parameterization of Nonlinear Model Predictive Control for Automotive Fuel Cell Systems

Thuc Anh Nguyen[1], Nikolai Weber[1] and Dirk Abel[1]

***Abstract*— This paper presents a nonlinear model predictive control strategy for automotive fuel cell system operation. The system-level air-path modeling approach, widely adopted in control-oriented studies, is extended by the critical aspect of membrane hydration. The control problem formulation reflects the task of dynamic power tracking and efficiency-optimized actuation of the peripheral components as well as adherence to system constraints to avoid harmful operation. This is combined with an analysis of the design parameters sampling time, integration scheme, and prediction horizon for efficient transcription of the optimal control problem. Closed-loop simulation results, conducted using a sampling rate of 8 ms, the standard fourth-order explicit Runge-Kutta method with one integration step and a horizon length of 25, successfully meet the control objectives.**

I. INTRODUCTION

Proton exchange membrane fuel cells (PEMFCs) convert chemical energy directly into electricity through the electrochemical reaction between hydrogen and oxygen, producing water and waste heat as byproducts. PEMFCs hold great promise for automotive applications as they offer several advantages: efficiencies of 50-55% surpassing those of internal combustion engines (15-35%), zero local emissions, superior power and energy density compared to rechargeable batteries, and the ability to refuel rapidly [1]. An automotive fuel cell system (FCS) comprises the fuel cell (FC) stack and various peripheral components: the air-path subsystem, hydrogen supply subsystem, water management subsystem, thermal management subsystem, and electrical subsystem. The air-path subsystem, which supplies oxygen by pressurizing ambient air using an electric compressor, holds particular significance for several reasons. Firstly, the compressor consumes up to 30% of the FC stack power output, making it a significant parasitic load. Secondly, the speed of power delivery by the FCS is limited by the dynamics of the air path, which are slower than those of the hydrogen supply and the electrochemical reaction. Thirdly, the air path is subject to critical constraints, including an oxygen excess ratio (OER) as well as compressor surge and choke [2]. In addition to the well-established issue of air-path control, the performance of the PEMFC itself depends on the proton conductivity of its membrane. Thus, balancing the membrane's water content between too dry conditions with high membrane resistance and too wet conditions, where liquid water blocks the reactant pathways, is essential [3].

[1]Thuc Anh Nguyen, Nikolai Weber and Dirk Abel are with the Institute of Automatic Control, RWTH Aachen University, 52074 Aachen, Germany thuc.anh.nguyen@rwth-aachen.de, n.weber@irt.rwth-aachen.de, d.abel@irt.rwth-aachen.de

Overall, the control objectives are (i) dynamic power response, (ii) high FCS efficiency, and (iii) preventing safety-critical and lifetime-limiting states. Particularly during dynamic operation and due to the different time scales, maintaining operational limits is challenging [3].

A. State of the Art

Various studies have applied model predictive control (MPC) to FCSs with a focus on air-path dynamics, e.g. [2], [4]. A few have broadened the system's scope by including other subsystems and providing more detailed FC modeling for a more realistic representation and the enforcement of crucial constraints outside the air path, as seen in [5], [6]. In all these MPC applications to FCSs, the characteristic nonlinear system dynamics are linearized around the current operating point (OP).

Nonlinear MPC (NMPC) directly optimizes the nonlinear system dynamics, potentially enhancing prediction accuracy, especially during transients. Luna et al. [7] apply NMPC to a FCS with a focus on degradation and other aspects beyond the air-path dynamics. Neisen et al. propose two NMPC variants for automotive FCSs in [8], [9] with a focus on the air-path dynamics, where [9] also considers the FCS in a hybrid setting with a battery system. They employ the `GRAMPC` toolbox [10] to solve the continuous-time optimal control problem (OCP). In contrast, this work uses the direct method. Direct methods offer advantages over indirect ones, as they handle inequality constraints more easily, with established approaches [11]. Schmitt et al. propose a tracking NMPC for the air-path control of FCSs tailored to a 10 kW test bench [12]. Despite the small power not being representative of the automotive application, they notably investigate aspects of real-time capability, including the choice of integration scheme and condensing method [13], and evaluate their control algorithms on embedded hardware tailored to automotive environments.

B. Contribution

This paper extends the work of Neisen et al. and Schmitt et al. application-wise by specifically considering the humidification state of the membrane in addition to the air path in the NMPC design. We combine this with numerical analyses to guide the selection of the design parameters sampling time, integration scheme, and prediction horizon, taking into account both computational cost and control performance.

The paper is structured as follows: In Section II, we present the control objectives and structure. Section III describes the dynamic equations. Section IV outlines the NMPC's OCP

Copyright ©2024 EUCA

formulation, and Section V covers the aforementioned numerical analyses for its efficient transcription into a nonlinear program (NLP). Finally, Section VI presents closed-loop simulation results.

II. CONTROL TASK & STRUCTURE

Given the current state of a FC dominant electric vehicle, which comprises a FCS and a small hybridization battery, the task of the control algorithm is to determine the optimal actuator inputs so that the vehicle's power output, P_{drive}, matches the power demand, all while maintaining safe and efficient operation. For the context of this work, this control task is divided between two controllers. The high-level NMPC incorporates a static model of the FCS and a dynamic model of the battery. Minimizing the FCS's efficiency and charge depletion of the battery, it determines the optimal power split between the two systems. The statically optimized operating points of the FCS can further serve as references for the peripheral components in the low-level NMPC, which directly actuates the fuel cell electric vehicle (FCEV). For that purpose, the low-level NMPC incorporates a dynamic model of the FCS and a static model of the battery. It tracks the references provided by the high-level controller and thus ensures that the power demand is dynamically met and FCS efficiency is statically maximized. At the same time, it avoids critical states which would lead to damage or aging of the FCS components. Here, we specifically consider preventing oxygen starvation, compressor surge and choke as well as maintaining proper membrane hydration.
The division of the control task between two controllers is motivated by the different relevant time scales, which are slower for the power-split optimization than for the low-level controller. The latter is focus of this work and its controller-internal model is presented in the following.

III. SYSTEM DYNAMICS

The control-oriented model of the FCEV primarily comprises a dynamic FCS model and a static battery model, as illustrated in Fig. 1. Details of the battery model, along

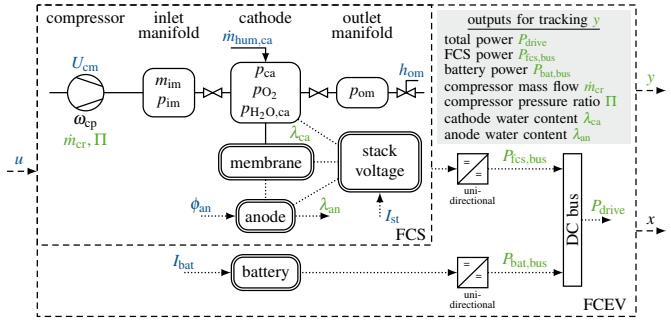

Fig. 1: Schematic overview of the model structure used for control-oriented modeling the FCEV.

with power loss considerations in the DCDC converters, are addressed in [14]. This section focuses on the FCS model, comprising the dynamics of the air path and the static relationships for stack voltage, membrane humidity, and anode. Our model is largely based on the zero-order model by [3], which is primarily developed using first principles. It is based on the following assumptions: (1) All gases are treated as ideal gases; (2) The temperature of the FC stack is perfectly controlled by the cooling system, ensuring a constant and uniform stack temperature of $T_{\text{st}} = 80\,^{\circ}\text{C}$; (3) The properties of the flow exiting the cathode (temperature, pressure, humidity) are assumed to be the same as those inside the cathode and are considered to dominate the reaction at the catalyst layers; (4) Water inside the cathode exists solely in vapor form. Once the relative humidity exceeds 100 %, it instantly condenses to liquid form and is removed. Thus, flooding is not modeled; (5) Finally, spatial variations are neglected.
Since the FCS model in [14], consisting of the air path and the stack voltage, is expanded here for the consideration of membrane humidification, this section concentrates on the submodels which are added to or modified from [14].

A. Cathode flow model

For the cathode airflow dynamics, we apply mass continuity and the ideal gas law to oxygen, nitrogen and water vapor inside the cathode volume V_{ca}. This yields the following differential equations for the partial pressure of oxygen, p_{O_2}, nitrogen, p_{N_2}, and vapor, $p_{\text{H}_2\text{O,ca}}$:

$$\frac{\mathrm{d}p_{\text{O}_2}}{\mathrm{d}t} = \frac{R \cdot T_{\text{st}}}{M_{\text{O}_2} \cdot V_{\text{ca}}} \cdot \left(\dot{m}^{\text{out}}_{\text{O}_2,\text{im}} - \dot{m}^{\text{react}}_{\text{O}_2} - \dot{m}^{\text{out}}_{\text{O}_2,\text{ca}}\right), \tag{1}$$

$$\frac{\mathrm{d}p_{\text{N}_2}}{\mathrm{d}t} = \frac{R \cdot T_{\text{st}}}{M_{\text{N}_2} \cdot V_{\text{ca}}} \cdot \left(\dot{m}^{\text{out}}_{\text{N}_2,\text{im}} - \dot{m}^{\text{out}}_{\text{N}_2,\text{ca}}\right), \tag{2}$$

$$\begin{aligned}\frac{\mathrm{d}p_{\text{H}_2\text{O,ca}}}{\mathrm{d}t} = \frac{R \cdot T_{\text{st}}}{M_{\text{H}_2\text{O}} \cdot V_{\text{ca}}} \Big(&\dot{m}^{\text{out}}_{\text{H}_2\text{O,im}} + \dot{m}^{\text{gen}}_{\text{H}_2\text{O}} + \dot{m}_{\text{H}_2\text{O,mem}} \\ &+ \dot{m}^{\text{hum,ca}}_{\text{H}_2\text{O}} - \dot{m}^{\text{out}}_{\text{H}_2\text{O,ca}} - \dot{m}^{\text{cond}}_{\text{H}_2\text{O}}\Big),\end{aligned} \tag{3}$$

where R is the universal gas constant and M_i the molar mass of the respective species $i \in \{\text{O}_2, \text{N}_2, \text{H}_2\text{O}\}$. The inlet and outlet mass flow rates, $\dot{m}^{\text{out}}_{i,\text{im}}$ and $\dot{m}^{\text{out}}_{i,\text{ca}}$ in (1) to (3), are calculated using thermodynamic properties. The former are:

$$\dot{m}^{\text{out}}_{\text{O}_2,\text{im}} = \frac{x_{\text{O}_2,\text{atm}}}{1 + w_{\text{im}}} \cdot \dot{m}^{\text{out}}_{\text{im}}, \tag{4}$$

$$\dot{m}^{\text{out}}_{\text{N}_2,\text{im}} = \frac{1 - x_{\text{O}_2,\text{atm}}}{1 + w_{\text{im}}} \cdot \dot{m}^{\text{out}}_{\text{im}}, \tag{5}$$

$$\dot{m}^{\text{out}}_{\text{H}_2\text{O,im}} = \frac{w_{\text{im}}}{1 + w_{\text{im}}} \cdot \dot{m}^{\text{out}}_{\text{im}}, \tag{6}$$

where $\dot{m}^{\text{out}}_{\text{im}}$ is the inlet manifold exit flow, $x_{\text{O}_2,\text{atm}}$ the molar fraction of oxygen in ambient air, and w_{im} the humidity ratio in the inlet manifold. The latter is calculated as follows

$$w_{\text{im}} = \frac{M_{\text{H}_2\text{O}}}{x_{\text{O}_2,\text{atm}} \cdot M_{\text{O}_2} + (1 - x_{\text{O}_2,\text{atm}}) \cdot M_{\text{N}_2}} \cdot \frac{\phi_{\text{atm}} \cdot p_{\text{sat}}}{p_{\text{atm}} - \phi_{\text{atm}} \cdot p_{\text{sat}}},$$

where p_{sat} is the saturation pressure at T_{st}, and p_{atm} and Φ_{atm} are the ambient pressure and relative humidity, respectively. A linearized nozzle equation is used to calculate the cathode exit flow rate, $\dot{m}_{\text{ca,out}}$:

$$\dot{m}_{\text{ca,out}} = k_{\text{ca,out}}(p_{\text{ca}} - p_{\text{om}}), \tag{7}$$

where $p_{ca} = p_{O_2} + p_{N_2} + p_{v,ca}$ is the cathode total pressure, p_{om} the outlet manifold pressure, and $k_{ca,out}$ the nozzle constant. With $\dot{m}_{ca,out}$ defined in (7), the mass flow rate of each species at the cathode exit, $\dot{m}^{out}_{i,ca}$, is given by:

$$\dot{m}^{out}_{i,ca} = \frac{p_i \cdot M_i}{p_{O_2} \cdot M_{O_2} + p_{N_2} \cdot M_{N_2} + p_v \cdot M_{H_2O}} \cdot \dot{m}^{out}_{ca} \quad (8)$$

Electrochemistry principles are used to calculate the rates of oxygen consumption, $\dot{m}^{react}_{O_2}$, and water vapor generation, $\dot{m}^{gen}_{H_2O}$, from the stack current I_{st}:

$$\dot{m}^{react}_{O_2} = \frac{M_{O_2} \cdot n \cdot I_{st}}{4 \cdot F}, \qquad \dot{m}^{gen}_{H_2O} = \frac{M_{H_2O} \cdot n \cdot I_{st}}{2 \cdot F}, \quad (9)$$

where F is the Faraday Constant and n the number of FCs in the stack. The net vapor mass flow across the membrane, $\dot{m}_{H_2O,mem}$, is calculated in the membrane hydration model, and the vapor injected by the humidifier, $\dot{m}^{hum,ca}_{H_2O}$, is a control input. Lastly, the amount of vapor that condenses instantly once $p_{H_2O,ca}$ exceeds p_{sat}, denoted as $\dot{m}^{cond}_{H_2O}$, is defined as

$$\dot{m}^{cond}_{H_2O} = \Psi \cdot k_{cond} \cdot \frac{M_{H_2O} \cdot V_{ca}}{R \cdot T_{st}} \cdot (p_{H_2O,ca} - p_{sat}), \quad (10)$$

where k_{cond} is the condensation rate constant and Ψ the transition function between condensation $(p_{H_2O,ca} > p_{sat})$ and no condensation $(p_{H_2O,ca} \leq p_{sat})$. To avoid hard switching between 0 and 1, we use an approximation with a sigmoid function [15]:

$$\Psi = \frac{1}{\pi} \cdot \arctan\left(k_{sig} \cdot [p_{v,ca} - p_{sat}]\right) + \frac{1}{2}, \quad (11)$$

The constant k_{sig} determines the width of the transition area, thus balancing approximation error and numerical stiffness.

B. Membrane hydration model

The membrane hydration model based on [3] calculates the water flow across the membrane and its water content. Both are assumed to be uniform across the membrane's surface. The water transport occurs via two distinct mechanisms. The first, electro-osmotic drag, involves water molecules being dragged from anode to cathode by the protons, where the amount of water is proportional to the electro-osmotic drag coefficient, n_d. Secondly, back-diffusion, driven by the gradient of water concentration across the membrane, typically moves water from cathode to anode, where we assume the water concentration to change linearly over the membrane thickness, t_{mem}. Combining the two mechanisms, the total molar flux from anode to cathode is

$$\dot{n}_{H_2O,mem,fc} = n_d \cdot \frac{I_{st}}{A_{fc} \cdot F} - D_w \cdot \frac{c_{H_2O,ca} - c_{H_2O,an}}{t_{mem}}, \quad (12)$$

where A_{fc} denotes the cell area and $c_{H_2O,i}$ with $i \in \{an, ca\}$ the water concentration at the anode and cathode. The coefficient n_d and the diffusion coefficient of water through the membrane D_w vary with the membrane water content λ_{mem}:

$$n_d = 0.0029 \cdot \lambda^2_{mem} + 0.05 \cdot \lambda_{mem} - 3.4 \cdot 10^{-19} \quad (13)$$

$$D_w = D_\lambda \cdot \exp\left(2416 \cdot \left(\frac{1}{303} - \frac{1}{T_{st}}\right)\right) \quad (14)$$

$$D_\lambda = \begin{cases} 10^{-6} & ,\lambda_{mem} < 2 \\ 10^{-6} \cdot (1 + 2 \cdot (\lambda_{mem} - 2)) & ,2 \leq \lambda_{mem} \leq 3 \\ 10^{-6} \cdot (3 - 1.67 \cdot (\lambda_{mem} - 3)) & ,3 < \lambda_{mem} < 4.5 \\ 1.25 \cdot 10^{-6} & ,\lambda_{mem} \geq 4.5 \end{cases}$$

The membrane water content, as well as the water content of the anode, λ_{an}, and cathode, λ_{ca}, are calculated from the respective water activities a_i, $i \in \{an, ca, mem\}$:

$$\lambda_i = \begin{cases} 0.043 + 17.81 \cdot a_i - 39.85 \cdot a_i^2 + 36.0 \cdot a_i^3 & ,0 < a_i \leq 1 \\ 14 + 1.4 \cdot (a_i - 1) & ,1 < a_i \leq 3, \end{cases}$$

with the water activities given by

$$a_i = \frac{p_{H_2O,i}}{p_{sat}}, i \in \{an, ca\}, \quad a_{mem} = \frac{1}{2} \cdot (a_{an} + a_{ca}) \quad (15)$$

The vapor concentrations are calculated as follows:

$$c_{H_2O,i} = \frac{\rho_{mem,dry} \cdot \lambda_i}{M_{mem,dry}}, \quad i \in \{an, ca\}, \quad (16)$$

where $\rho_{mem,dry}$ denotes the membrane's dry density and $M_{mem,dry}$ the membrane dry equivalent weight. Finally, the total vapor mass flow considered for the stack is

$$\dot{m}_{H_2O,mem} = M_{H_2O} \cdot A_{fc} \cdot n \cdot \dot{n}_{H_2O,mem,fc}. \quad (17)$$

Note that the expressions for D_λ and λ_i violate the requirement for continuous differentiability. Therefore, the hard switching is again numerically smoothed similar to (11).

C. Other static relations

We do not consider a dynamic model of the anode. Instead, we assume perfect control of anode pressure p_{an}, i.e. no pressure difference across the membrane: $p_{an} = p_{ca}$. We further assume the relative humidity in the anode, Φ_{an}, to be a control input. The hydrogen partial pressure, p_{H_2}, is then given by

$$p_{H_2} = p_{an} - \phi_{an} \cdot p_{sat}. \quad (18)$$

Besides, in addition to the power consumed by the compressor motor, P_{cm}, the total stack power, P_{st}, is further diminished by the power consumption of the humidification subsystem, $P_{hum,ca}$ and $P_{hum,an}$. The constant term P_{aux} simply combines the power consumption of all remaining actuators.

$$P_{fcs,net} = P_{st} - P_{cm} - P_{hum,ca} - P_{hum,an} - P_{aux} \quad (19)$$

IV. CONTROL PROBLEM FORMULATION

A. Prediction model

The differential and algebraic equations that the controller-internal model comprises are generally referred to as:

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), p(t)) \\ y(t) &= g(x(t), u(t), p(t)), \end{aligned} \quad (20)$$

where $x \in \mathbb{R}^7$ and $u \in \mathbb{R}^6$ are the state and input vector, respectively, and $p \in \mathbb{R}^2$ is the run-time parameter vector, consisting of T_{st} and the battery state-of-charge. For our NMPC design, we adopt the delta formulation. This allows us to constrain the rates of change of the inputs while maintaining a sparse OCP structure of the resulting NLP,

given that multiple shooting is used. The approach extends the state vector, denoted as $\tilde{x}$ in the augmented state space, to include the actuated variables, while the augmented input vector $\tilde{u}$ comprises their rates of change:

$$\tilde{x} = \begin{bmatrix} p_{\mathrm{im}} & p_{\mathrm{O_2}} & p_{\mathrm{N_2}} & p_{\mathrm{H_2O,ca}} & p_{\mathrm{om}} & m_{\mathrm{im}} & \omega_{\mathrm{cp}} \\ & I_{\mathrm{st}} & U_{\mathrm{cm}} & h_{\mathrm{om}} & \dot{m}_{\mathrm{hum,ca}} & \Phi_{\mathrm{an}} & I_{\mathrm{bat}} \end{bmatrix}^\top \in \mathbb{R}^{13}$$

$$\tilde{u} = \begin{bmatrix} \dot{I}_{\mathrm{st}} & \dot{U}_{\mathrm{cm}} & \dot{h}_{\mathrm{om}} & \ddot{m}_{\mathrm{hum,ca}} & \dot{\Phi}_{\mathrm{an}} & \dot{I}_{\mathrm{bat}} \end{bmatrix}^\top \in \mathbb{R}^6$$

The outputs y we select for tracking are:

$$y = \begin{bmatrix} P^{\mathrm{drive}} & P^{\mathrm{fcs,bus}} & P^{\mathrm{bat,bus}} & \dot{m}_{\mathrm{cr}} & \Pi & \lambda_{\mathrm{ca}} & \lambda_{\mathrm{an}} \end{bmatrix}^\top \in \mathbb{R}^7$$

Note that, from this point forward, the augmented state $\tilde{x}$ and input vector $\tilde{u}$ will simply be referred to as x and u.

B. Optimal Control Problem

The NMPC is defined by the OCP in (21). The objective function regulates the outputs to their reference values $y_{\mathrm{ref}}(t) \in \mathbb{R}^7$ starting from the initial state $\hat{x}_0$ and penalizes input changes over the prediction time t_{p}. The last two terms in the cost functional penalize the violation of soft constraints, where $\varepsilon \in \mathbb{R}^9$ is the vector of slack variables.

$$\begin{aligned} \min_{x(\cdot),u(\cdot),\varepsilon} \quad & \frac{1}{2}\int_0^{t_{\mathrm{p}}} \left(\|y(t) - y_{\mathrm{ref}}(t)\|_Q^2 + \|u(t)\|_R^2 \right) \mathrm{d}t \\ & + \|\varepsilon\|_{S_{\mathrm{q}}}^2 + |\varepsilon|_{S_{\mathrm{l}}} \\ \text{s.t.} \quad & 0 = x(0) - \hat{x}_0, \\ & \dot{x}(t) = f(x(t), u(t), p(t)) \qquad \forall t \in [0, t_{\mathrm{p}}], \\ & 0 \geq h(x(t), u(t), p(t), \varepsilon) \quad \forall t \in [0, t_{\mathrm{p}}]. \end{aligned} \tag{21}$$

The inequality function h forces all states and actuators except for the battery current I_{bat} to be non-negative. Additionally, it considers actuator limits by imposing the following input constraints:

$$I_{\mathrm{st}}^{\min} \leq I_{\mathrm{st}}, \qquad h_{\mathrm{om}} \leq 1, \qquad \Phi_{\mathrm{an}} \leq 1, \tag{22}$$

where h_{om} is the opening position of the outlet manifold valve. The lower bound on the stack current ensures operation away from the open-circuit voltage, which is a catalyst and membrane stressor [5]. To accommodate actuator dynamics that are not explicitly modeled, we also constrain the rates of input changes, recognizing that this notably simplifies the typically more complex behavior of actuators.

$$u_{\min} \leq u \leq u_{\max} \tag{23}$$

The above constraints on the actuators are all enforced as hard constraints, i.e., no constraint violation is allowed.
In contrast, the safety and degradation constraints considered in this work are imposed as soft constraints to ensure feasibility of the problem. Safe operation of the compressor is guaranteed within the choke and surge boundary:

$$\begin{aligned} 13.125 \cdot \dot{m}_{\mathrm{cr}} + 0.82 - \Pi &\leq \varepsilon_{\mathrm{choke}} \\ \Pi - 278.6852 \cdot \dot{m}_{\mathrm{cr}}^2 - 17.8817 \cdot \dot{m}_{\mathrm{cr}} - 0.7888 &\leq \varepsilon_{\mathrm{surge}} \end{aligned} \tag{24}$$

Moreover, to ensure sufficient delivery of oxygen to the cathode channels, the OER $\lambda_{\mathrm{O_2}} = \dot{m}_{\mathrm{O_2,im}}^{\mathrm{out}} / \dot{m}_{\mathrm{O_2}}^{\mathrm{react}}$ is constrained:

$$\lambda_{\mathrm{O_2}} \leq \lambda_{\mathrm{O_2}}^{\min} + \varepsilon_{\lambda_{\mathrm{O_2}}} \tag{25}$$

Lastly, the water content in the anode, cathode and membrane, λ_i with $i \in \{\mathrm{ca}, \mathrm{an}, \mathrm{mem}\}$, is constrained:

$$\lambda_{i,\min} - \varepsilon_{\lambda_i,\min} \leq \lambda_i \leq \lambda_{i,\max} + \varepsilon_{\lambda_i,\max} \tag{26}$$

V. ANALYSES FOR PROBLEM TRANSCRIPTION

Using the so-called direct method, the continuous-time OCP in (21) is discretized. Here, multiple-shooting is applied and the resulting NLP reads:

$$\begin{aligned} \min_{\substack{x_0,\ldots,x_{N-1},\\ u_0,\ldots,u_{N-1},\\ \varepsilon}} \quad & \frac{1}{2}\sum_{k=0}^{N-1} \|y_k - y_{\mathrm{ref},k}\|_Q^2 + \|u_k\|_R^2 + \|\varepsilon\|_{S_{\mathrm{q}}}^2 + |\varepsilon|_{S_{\mathrm{l}}} \\ \text{s.t.} \quad & 0 = x_0 - \hat{x}_0, \\ & x_{k+1} = F(x_k, u_k, p_k) \quad k = 0, \ldots, N-1, \\ & 0 \geq h(x_k, u_k, p_k, \varepsilon) \; k = 0, \ldots, N-1 \end{aligned} \tag{27}$$

When transcribing the OCP, several design aspects influence the numerical size of the NLP, which affects both the computational cost of solving it and the resulting control performance. In this work, we analyze the choice of prediction horizon N, integration scheme, and sampling time T_{s}.

A. Sampling time T_{s}

The sampling time determines the upper bound of achievable control bandwidth. Ideally, we seek the control response time to be as fast as the fastest time constant of the prediction model. To analyze its time constants, we simulate the model with inputs corresponding to a power trajectory covering 100 different OPs from 10 kW to 50 kW. At each OP, the system dynamics are linearized and the eigenvalues λ_{OP} of the system matrix $A_{\mathrm{OP}} = \partial f / \partial x \big|_{\mathrm{OP}}$ are evaluated. We assume, that the reciprocal of each eigenvalue is a good representation of the autonomous response time of the corresponding system mode at the respective OP. We further assume that the resulting eigenvalues λ_{OP} are a good approximation of the dynamic behavior of the controller-internal model, since the simulation trajectory covers a wide range of different OPs.
Upon evaluating the time constants, we find the fastest system mode to have a time constant between 7.6 ms and 9 ms, while the system's slowest time constant is 1.3 s. The former provides a guide in selecting an appropriate order of magnitude for the sampling time. Here, we choose $T_{\mathrm{s}} = 8$ ms.

B. Numerical integration scheme and step size

Every numerical integration scheme applied to the prediction model corresponds to a discretization error and a computational cost, both of which we aim to minimize. To select an appropriate scheme, we investigate accuracy and computational cost for explicit Runge-Kutta (ERK) and implicit Runge-Kutta (IRK) methods with different stages s. For the latter, we utilize Gauss-Legendre (GL) collocation methods of order $p = 2 \cdot s$ and Radau IIa (RDIIa) methods of order $p = 2 \cdot s - 1$. This approach is adopted from [12]. All mentioned integration schemes are implemented in `MATLAB` using `acados` [16] and each scheme is referred to with its abbreviation and the number of stages s, i.e. ERK with four stages is denoted as ERK4.

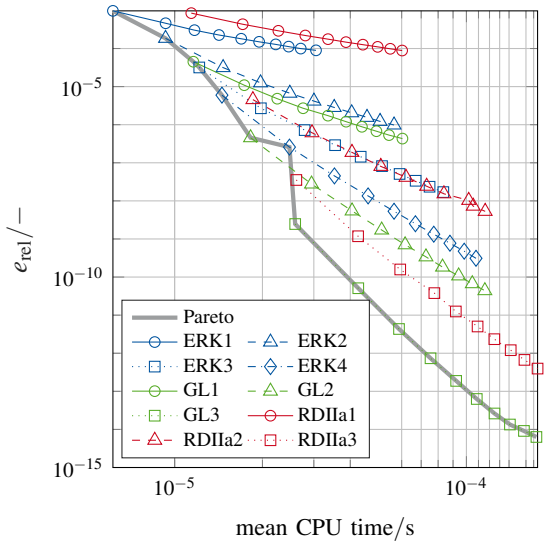

Fig. 2: Relative discretization error vs. mean CPU time for tested integration schemes.

As a measure for accuracy, we define the relative discretization error e_{rel}, for a simulation of N_{t} time steps starting from the initial state $\hat{x}_0$ with a constant input u, as follows:

$$e_{\mathrm{rel}} = \frac{1}{n_{\mathrm{x}}} \sum_{i=1}^{n_{\mathrm{x}}} \frac{1}{N_{\mathrm{t}}} \sum_{k=0}^{N_{\mathrm{t}}} \frac{|F(x_k,u_k,p_k) - F_{\mathrm{gt}}(x_k,u_k,p_k)|}{F_{\mathrm{gt}}(x_k,u_k,p_k)}, \tag{28}$$

where n_{x} is the number of states. The subscript "gt" refers to the ground truth discretization, which is numerically calculated with GL3 and 400 integration steps. We evaluate e_{rel} for each integration scheme and a range of integration steps from 1 to 10, where one integration step corresponds to a step size of 8 ms. Additionally, we determine the corresponding mean CPU time out of 4000 simulation repetitions, conducted on a Lenovo T480s laptop, with a Intel Core i7-8550U at 1.8 GHz with 16 GB RAM. The resulting relationship between e_{rel} and mean CPU time is shown in Fig. 2.

Depending on the required accuracy, different schemes provide the most computationally efficient approximation of the state trajectories, as illustrated by the Pareto front in Fig. 2. For accuracies of $e_{\mathrm{rel}} > 10^{-6}$, the integrations schemes ERK1, ERK2, GL1 and ERK4, each with one integration step, dominate, ranking from least to most accurate. For accuracies below $e_{\mathrm{rel}} = 10^{-6}$, mostly the IRK method GL dominates on the Pareto front, except for ERK4 with two integration steps. For high accuracies, GL3 with the highest order tested is the most efficient. For a desired accuracy in the vicinity of $e_{\mathrm{rel}} = 10^{-6}$, i.e. between ERK4 and GL2, each with one integration step, we opt for the less accurate but more efficient ERK4.

C. Prediction horizon N

Ideally, the choice of prediction time $t_{\mathrm{p}} = T_{\mathrm{s}} \cdot N$ ensures that the MPC can sufficiently predict how the manipulated variables may affect the cost or outputs of interest. In this work, we do not consider future references to be

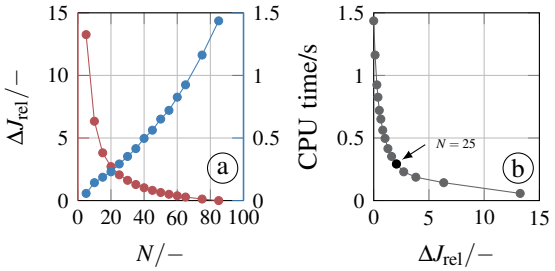

Fig. 3: (a) Relative cost difference and CPU time vs. prediction horizon, (b) CPU time vs. relative cost difference.

available in the prediction horizon. Thus, for the tracking of constant references, the slowest time constant of the prediction model may serve as a benchmark. However, we consider dynamically changing references with a time scale of one second here. At the same time, large horizons lead to more decision variables and constraints, i.e. to a larger optimization problem with longer execution times and higher memory footprint. Fig. 3a (blue) shows the evaluation of the average computation time for one control update in a closed-loop simulation conducted with a sampling time of $T_{\mathrm{s}} = 8$ ms and varying prediction horizons. Additionally, we examine the relative cost difference ΔJ_{rel} for each N as defined in (29), where $J(N) = \sum_{i=1}^{n_{\mathrm{c}}} J_i(N)/N$ is the sum over all control updates n_{c} of the average cost for one prediction interval at control update i. Among the prediction horizons tested, $J(N = 85)$ has the smallest value and is therefore used as the reference J_{ref}.

$$\Delta J_{\mathrm{rel}}(N) = \frac{J(N) - J_{\mathrm{ref}}}{J_{\mathrm{ref}}} \tag{29}$$

Fig. 3a (red) shows that the smaller N, the larger ΔJ_{rel}. Overall, the CPU time increases with larger N with a linear trend, while the cost improvement decreases with larger N with an exponential trend. Plotting these two measures against each other, we observe a trade-off curve, depicted in Fig. 3b, guiding the control designer to balance between computation time and cost. Here, we choose $N = 25$.

VI. SIMULATION RESULTS

We test the proposed control algorithm in an ideal closed-loop setting with no model mismatch. The implementation was carried out in `MATLAB`, using `CasADi 3.6.4` [17] with the solvers `IPOPT 3.14.11` and `MUMPS 5.4.1`. The NMPC is evaluated in a challenging scenario with rapid and substantial changes in power demand, see Fig. 4. The system trajectories subject to critical operational constraints are given in Fig. 5. For the high and sudden power demand at $t = 6$ s, we observe that the FCS power output struggles to meet its reference rapidly. This is due to both the lower limit on OER (Fig. 5b) and the compressor choke boundary (Fig. 5a), which coincide with the system trajectories during the power increase. This observation can be attributed to the slow dynamics of the air path. For this reason, the battery power overshoots its reference (Fig. 4c), so that the overall

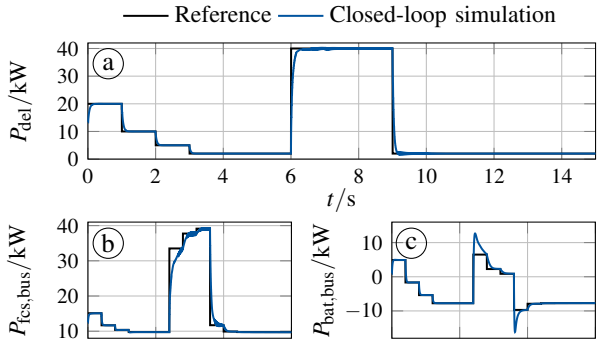

Fig. 4: Power reference tracking for (a) total power, (b) FCS power, (c) battery power.

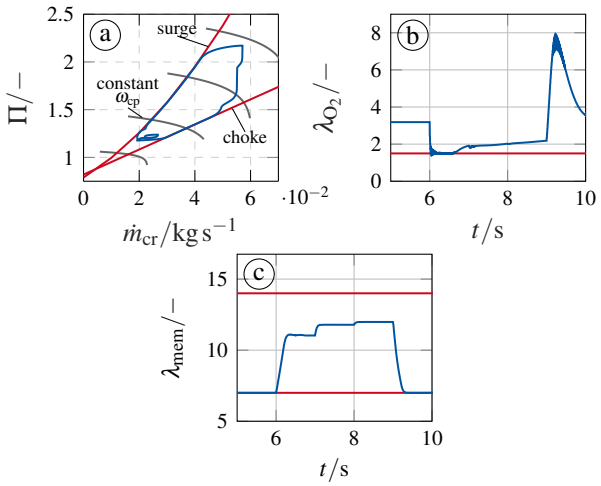

Fig. 5: (a) Compressor choke and surge boundaries, (b) OER constraint, (c) membrane water content limits.

power can be delivered dynamically (Fig. 4a). An equivalent observation can be made during the sudden and substantial decrease in power at $t = 9\,\mathrm{s}$, except that, there is no upper limit on the OER, and the limitation arises from compressor surge. The membrane water content, depicted in Fig. 5c, is kept within its bounds. When operating at a lower reaction rate, the membrane is minimally humidified to ensure it remains within its lowest permissible boundary. At higher reaction rates between 6 s and 9 s, λ_{mem} would automatically satisfy the constraints due to increased water production. Still, we observe actuation by the humidification system, resulting in higher λ_{mem} and slightly higher FCS efficiency.

VII. Conclusions and Future Work

We have introduced an NMPC approach to manage dynamic power delivery with consideration of the air-path and water management subsystem in automotive FCSs. Our study examines the design parameters sampling time, integration scheme, and prediction horizon to find a balance between control effectiveness and computational efficiency. In a demanding evaluation scenario, the proposed controller meets the set control objectives and adheres to critical constraints, laying a groundwork for further investigation into water management's effect on the overall system performance. Future efforts will focus on employing advanced numerical optimal control techniques to develop a real-time capable controller. Additionally, we aim to refine the prediction model by incorporating factors like liquid water dynamics and other relevant effects.

References

[1] O. Z. Sharaf and M. F. Orhan, "An overview of fuel cell technology: Fundamentals and applications," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 810–853, 2014.

[2] A. Vahidi, A. G. Stefanopoulou, and H. Peng, "Current management in a hybrid fuel cell power system: A model-predictive control approach," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 6, pp. 1047–1057, 2006.

[3] J. T. Pukrushpan, A. G. Stefanopoulou, and H. Peng, *Control of Fuel Cell Power Systems: Principles, Modeling, Analysis and Feedback Design*. Springer Science & Business Media, 2004.

[4] S. Hahn, J. Braun, H. Kemmer, and H.-C. Reuss, "Adaptive operation strategy of a polymer electrolyte membrane fuel cell air system based on model predictive control," *International Journal of Hydrogen Energy*, vol. 46, no. 33, pp. 17 306–17 321, 2021.

[5] A. Goshtasbi and T. Ersal, "Degradation-conscious control for enhanced lifetime of automotive polymer electrolyte membrane fuel cells," *Journal of Power Sources*, vol. 457, p. 227996, 2020.

[6] M. Vrlić, D. Ritzberger, and S. Jakubek, "Safe and efficient polymer electrolyte membrane fuel cell control using successive linearization based model predictive control validated on real vehicle data," *Energies*, vol. 13, no. 20, p. 5353, 2020.

[7] J. Luna, S. Jemei, N. Yousfi-Steiner, A. Husar, M. Serra, and D. Hissel, "Nonlinear predictive control for durability enhancement and efficiency improvement in a fuel cell power system," *Journal of Power Sources*, vol. 328, pp. 250–261, 2016.

[8] V. Neisen, J. Mannhardt, and D. Abel, "Dynamic tracking of power demand for integrated fuel cell systems using nonlinear model predictive control," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 13 216–13 223, 2020.

[9] D. Klöser, "Hierarchical model predictive control for the dynamical power split of a fuel cell hybrid vehicle," Master's thesis, RWTH Aachen University. Institute of Automatic Control, 2019.

[10] T. Englert, A. Völz, F. Mesmer, S. Rhein, and K. Graichen, "A software framework for embedded nonlinear model predictive control using a gradient-based augmented lagrangian approach (grampc)," *Optimization and Engineering*, vol. 20, no. 3, pp. 769–809, 2019.

[11] B. Käpernick, "Gradient-based nonlinear model predictive control with constraint transformation for fast dynamical systems," Ph.D. dissertation, Universität Ulm.

[12] L. Schmitt and D. Abel, "Numerical integration for nonlinear model predictive control of a fuel cell system," in *2023 American Control Conference (ACC)*. IEEE, 2023.

[13] L. Schmitt, N. Nickig, M. Bahr, S. Gößling, and D. Abel, "Review, evaluation and application of condensing algorithms for model predictive control based on a first-order method," in *2023 European Control Conference (ECC)*. IEEE, 2023.

[14] T. A. Nguyen, V. Neisen, and D. Abel, "Investigation of the real-time feasibility of nmpc for air-path control in automotive fuel cell systems," in *AmE 2024 -Automotive meets Electronics; 14. GMM Symposium*, pp. 67–72.

[15] D. Ritzberger, J. Höflinger, Z. P. Du, C. Hametner, and S. Jakubek, "Data-driven parameterization of polymer electrolyte membrane fuel cell models via simultaneous local linear structured state space identification," *International Journal of Hydrogen Energy*, vol. 46, no. 21, pp. 11 878–11 893, 2021.

[16] R. Verschueren, G. Frison, D. Kouzoupis, J. Frey, N. van Duijkeren, A. Zanelli, B. Novoselnik, T. Albin, R. Quirynen, and M. Diehl, "acados—a modular open-source framework for fast embedded optimal control," *Mathematical Programming Computation*, vol. 14, no. 1, pp. 147–183, 2022.

[17] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.