Energy Reports 10 (2023) 1–14

Contents lists available at ScienceDirect

Energy Reports

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

Research paper

Reduced-dimensionality nonlinear distributed-parameter observer for fuel cell systems

Martin Vrlić [a,*], Dominik Pernsteiner [a,b], Alexander Schirrer [a], Christoph Hametner [a,b], Stefan Jakubek [a]

[a] Institute of Mechanics and Mechatronics, TU Wien, Austria
[b] CD Laboratory for Innovative Control and Monitoring of Automotive Powertrain Systems, TU Wien, Austria

ARTICLE INFO

Article history:
Received 29 January 2023
Received in revised form 31 May 2023
Accepted 2 June 2023
Available online 15 June 2023

Keywords:
Fuel cell
Balanced truncation
Distributed state estimation
Extended kalman filter

ABSTRACT

To ensure reliable and efficient operation of fuel cell systems, it is important to monitor them online. However, placing sensors inside the fuel cell is often challenging, so virtual sensing using an efficient state observer is used in this study. Detecting local internal phenomena, such as reactants' starvation, membrane dryout/flooding, and nitrogen accumulation, requires knowledge of the spatial distribution of internal states. Lumped-parameter models are not suitable for this, as they use a single variable to describe parameters such as hydrogen concentration. Instead, a high-order distributed-parameter fuel cell model is used to predict the spatial profiles of various internal states. An observer algorithm is employed to correct the predicted quantities using a few measurements taken at the system boundary. This update step only considers dominant dynamics from a reduced model to adjust all system states accordingly, making it computationally efficient and robust. The observer algorithm's performance was verified against a high-fidelity model through detailed simulations.

1. Introduction

The share of hydrogen in the overall energy sector is expected to increase due to the transition to renewable energy sources. Fuel cells play an important role in converting hydrogen back into power. The application of fuel cells is envisaged both in transport (cars, trucks, railways, and aviation) and in stationary systems (industry and buildings) (IEA, 2019, 2021).

Polymer electrolyte membrane fuel cells (PEMFC) are an attractive zero-emission power source due to their high efficiency and other benefits such as low operating temperature, and fast startup (İnci et al., 2021). Fuel cells in automotive applications are of particular interest, and their development is a challenging and wide research area. One of the major topics still not fully resolved and under investigation is onboard system monitoring and diagnostics of fuel cell electric vehicles (FCEV). Evidently, good system observation is crucial for safe and efficient operation under highly dynamic conditions which are usually present in FCEVs. Measurement signals commonly available on FCEVs such as system voltage, anode/cathode inlet/outlet pressures and humidity are of course necessary, but not sufficient to detect critical conditions that may occur inside the system during operation. Information about internal states such as species concentrations in gas channels (GC) and gas diffusion layers (GDL), species densities and membrane water content, to name a few, gives more profound insight necessary for proper system monitoring and diagnostics. Methods for measuring the internal states of the system have been developed in the past. For example, neutron imaging and magnetic resonance imaging have been commonly used to measure water concentration inside a fuel cell (Wang et al., 2021; Lee et al., 2021). Furthermore, nitrogen and oxygen concentrations have been measured using gas chromatography analysis (Dobrokhotov and Larin, 2019). Temperature at certain points inside the cell has been determined using distributed fibre optic sensors (Zaghloul et al., 2021). However, measuring devices for such quantities are expensive, relatively big, hard to implement even under laboratory conditions, and outright impossible in vehicles. These methods have been summarised in Table 1.

A promising alternative, which is explored in this paper, is virtual sensing using a state observer. The state observer uses a model and uses available measurement signals to estimate the internal states of the system. An observer generally works as follows: a model is used to predict the same signals that are being measured by the available sensors, called outputs. In doing so, the internal states are estimated. The states are controlled using an algorithm to match the measured outputs. When the outputs are matched and if the system is observable, the internal states of the model match the unmeasurable reality. It is to be noted that for this to hold, the plant-model mismatch must be minimal.

* Corresponding author.
E-mail address: martin.vrlic@tuwien.ac.at (M. Vrlić).

https://doi.org/10.1016/j.egyr.2023.06.006

Nomenclature

State variables of the fuel cell quasi-2D model

γ_{H_2}	Hydrogen mole fraction
γ_{N_2}	Nitrogen mole fraction
γ_{O_2}	Oxygen mole fraction
λ	Membrane water content
ρ	Gas density
ξ_{H_2}	Hydrogen mass fraction
ξ_{N_2}	Nitrogen mass fraction
ξ_{O_2}	Oxygen mass fraction
i	Local current density
p	Gas pressure
V	Cell voltage
v	Gas velocity

Number of discretisation nodes

N_{ca}	Number of discretisation nodes in the anode channel
N_{ca}	Number of slices
N_{cc}	Number of discretisation nodes in the cathode channel
N_{ga}	Number of discretisation nodes in the anode gas diffusion layer
N_{gc}	Number of discretisation nodes in the cathode gas diffusion layer
N_{mem}	Number of discretisation nodes in the membrane

State space model

$\boldsymbol{\chi}$	State vector of the whole system
$\boldsymbol{\chi}_{ca}$	State vector of the anode channel
$\boldsymbol{\chi}_{cc}$	State vector of the cathode channel
$\boldsymbol{\chi}_{ga}$	State vector of the anode gas diffusion layer
$\boldsymbol{\chi}_{gc}$	State vector of the cathode gas diffusion layer
$\boldsymbol{\chi}_{mem}$	State vector of the membrane
$\boldsymbol{\mathcal{A}}$	System matrix
$\mathfrak{b}$	Right-hand side vector
$\boldsymbol{u}$	Input vector
$\boldsymbol{x}$	State vector of the whole system including past states
$\boldsymbol{y}$	Output vector
k	Current time step

Observer

$\tilde{\boldsymbol{Q}}$	Reduced order process noise covariance
$\boldsymbol{Q}$	Process noise covariance
$\boldsymbol{R}$	Measurement noise covariance
$\Delta\hat{\boldsymbol{x}}$	State correction
$\hat{\boldsymbol{x}}^-$	Predicted state vector
$\hat{\boldsymbol{y}}^-$	Predicted output vector
$\boldsymbol{\mathcal{T}}$	Transformation matrix from full state-space to reduced state-space
$\boldsymbol{y}_{meas}$	Measurement vector
$\tilde{\boldsymbol{A}}_r$	Reduced order Jacobian matrix
$\tilde{\boldsymbol{C}}_r$	Reduced order output matrix
$\tilde{\boldsymbol{K}}$	Kalman gain
$\tilde{\boldsymbol{P}}$	Updated error covariance matrix
$\tilde{\boldsymbol{P}}^-$	Predicted error covariance matrix
$\tilde{\boldsymbol{S}}$	Innovation covariance

Balanced truncation

$\boldsymbol{x}_s$	Scaled state vector
$\tilde{\boldsymbol{T}}$	Transformation matrix from physical state-space to balanced realisation
$\tilde{\boldsymbol{x}}$	State vector in the balanced realisation
$\tilde{n}_r$	Number of modes
$\tilde{\boldsymbol{x}}_e$	Eliminated states
$\tilde{\boldsymbol{x}}_r$	State vector in the reduced state-space
n_x	Number of states

Therefore, before the model can be used as a state observer, it must first be validated. In this work, however, as it is explained later on, the simulated reality and the model used as basis for the observer have the same underlying structure, but different number of discretisation points. For that reason, the validation of the model used by the observer is considered to be fulfilled.

Fuel cell models which are appropriate for observer design can be categorised on the basis of their underlying system dynamics (Hidayat et al., 2011) into:

- Lumped-parameter models, whose dynamics are described by ordinary differential equations.
- Distributed-parameter models, whose dynamics are described by partial differential equations.

The vast majority of fuel cell system observers were developed with lumped-parameter models with aggregated states. However, pressure, gas concentration etc. are non-uniformly distributed inside the cell. Distributed-parameter models are, therefore, necessary to describe along-the-channel effects and local phenomena.

Literature on observers for distributed states in fuel cells is scarce, as evident from a recent review on the topic of fuel cell observers (Yuan et al., 2020). A simplified, 10-element model for observing the water distribution profile is reported in Sarmiento-Carnevali et al. (2017). However, the focus of that paper is developing the controller. Concerning publications with a detailed description of the observer development, to the best of our knowledge, only Luna et al. tackled this topic in their work (Luna et al., 2015, 2016b,a, 2017). Their fuel cell observer employs a finite volume model and is based on assumed measurements of species concentration at the system boundary. In their paper, they only show the time-evolution of the species at the middle point of the channels.

In their work they use a sliding mode observer, a discontinuous type of observer that uses some input variable as corrective action. This input variable does not have a clear meaning and it is not seen in the results section of the paper, making it difficult to grasp the overall result. The complexity of an observer, such as the one proposed by Luna et al. naturally increases with the number of discretisation volumes (system resolution). This problem is resolved in the present work by using dominant-modes-only Jacobians for the update step of the observer (explained later on), resulting in an efficient observer algorithm. The observer is based on easily attainable measurements at the system boundary such as the cell voltage as a measured output and other commonly known quantities, namely the inlet/outlet pressures, the inlet gas composition, and the current density demand as boundary conditions. The actual distribution of the internal states of the system is presented in the figures in Section 6. The time evolution

Table 1
Measurement techniques for fuel cell internal states.

Ref	Measurement	Method
Wang et al. (2021), Lee et al. (2021)	Water concentration	Neutron imaging, magnetic resonance
Dobrokhotov and Larin (2019)	N_2 and O_2 concentrations	Gas chromatography
Zaghloul et al. (2021)	Local temperature	Distributed fibre optic sensors

Table 2
Existing observers for estimation of the internal states of fuel cell systems.

Ref	Observer type	Model type	Performance object(s)
Yuan et al. (2020) (Review paper)	Extended Kalman filter	Single phase, 8th order, lumped	Oxygen partial pressure
	Unscented Kalman filter	Single phase, 4th order, lumped	Nitrogen partial pressure at anode
	Luenberger	Two-phase, 6th order, lumped	Current density difference
	First-order sliding mode	Single phase, 6th order, lumped	Manifold pressure and mass flow
	High order sliding mode	Two-phase, 13th order, lumped	Hydrogen and oxygen partial pressures
Luna et al. (2015, 2016b,a, 2017)	High order sliding mode	Single phase, distributed	Hydrogen and oxygen partial pressures along the channel

Table 3
Reduction techniques for Extended Kalman Filters with high-order models as its basis.

Ref	Observer purpose	Reduction approach
Farrell and Ioannou (2001)	Storm track forecasting	Balanced truncation
Khodadadi and Jazayeri-Rad (2011)	Joint state and parameter estimation of a continuous stirred tank reactor	Dual EKF
Lee et al. (2007)	Li-ion battery SOC estimation	Equivalent circuit model simplification
Park et al. (2013)	Estimation of internal states of permanent magnet synchronous motors	Parallel structure of simplified models
Pernsteiner et al. (2021)	Temperature estimation of a latent heat storage	Balanced truncation

of the whole distribution is shown in the videos found in the supplementary material. In the present paper, a detailed quasi-2D fuel cell model from Murschenhofer et al. (2018) is used as the basis for a newly developed distributed-parameter fuel cell observer. The representative observer types have been collected in Table 2.

The observer algorithm chosen in this work is based on the well-known extended Kalman filter (EKF) (Jazwinski, 1970; Wishner et al., 1969).

When applying an EKF to high-order models, such as the discretised quasi-2D fuel cell model, two major problems can arise: (1) the problem becomes infeasible to be solved in real time due to its computational complexity, and (2) the system becomes (almost) unobservable when the entire state dimension is considered. To overcome these challenges, different types of reduced-order EKFs have been investigated, see for example Farrell and Ioannou (2001), Khodadadi and Jazayeri-Rad (2011), Lee et al. (2007), Park et al. (2013), Pernsteiner et al. (2021). Model reduction methods in observer design can be used to compute reduced-order system Jacobians necessary for the update step in the EKF, explained in Section 4. In Table 3, different methods of dealing with the problem of high-dimensionality are shown.

Model reduction approaches are applied to efficiently reduce the system dynamics' complexity while maintaining their dominant behaviour. Especially model reduction in linear systems is well established, and the research topic has been broadly covered in literature, see for example the review paper (Benner et al., 2015). An example for linear model reduction is the balanced truncation introduced by Moore (1981) where the modes of the system showing the smallest Hankel singular values are eliminated. For nonlinear systems, data-based approaches such as the proper orthogonal decomposition (Brunton and Kutz, 2019) or the dynamic mode decomposition (Schmid, 2010) can be an appropriate choice. The main idea behind these approaches is the empirical investigation of data representative for a relevant operating range. The data are decomposed into orthonormal basis vectors using singular value decomposition (SVD). Few basis vectors that describe the most characteristic system behaviour can be extracted, and the system properties can be projected onto these so-called *modes*.

In this work, the model reduction technique applied to obtain the reduced-order system Jacobians is Balanced truncation. Only the most dominant modes, whose Hankel singular values exceed a defined threshold, are chosen to be reflected in the Jacobians, the others are discarded. Particularly, the system is being successively linearised, ensuring the validity of the Jacobian matrix used in the observer. A simulation study is performed to test the aforementioned approach. Measurements are taken at the system boundary with the goal of estimating the unmeasurable internal states. The full order quasi-2D model is used for prediction and validation, but these two versions of the model differ in grid sizes. Despite this model mismatch, the observer algorithm works well. For the update step of the EKF, the reduced order system Jacobians are used. The described procedure is depicted in Fig. 1. All details and results are discussed in Section 6. The identified research gap is to detect the distributed internal states of a fuel cell in an efficient way with realistic measurements. The fuel cell observer proposed for this purpose, which is based on a high-fidelity model with distributed parameters and a dominant-modes-only correction, has not yet been reported in the fuel cell community. In particular, the temporal evolution of the internal states' spatial distribution is shown, highlighting the extent of such an observer's capabilities.

In summary, this paper shows an internal state estimation framework using an on-board commonly available measurement signal, i.e. system voltage. Moreover, the estimated states are not lumped as in the majority of the literature, but their distribution is shown. In literature (Luna et al., 2015), crucial simplifying assumptions were made which were not made in this paper. These include: no nitrogen crossover, availability of outlet concentrations measurements and assuming the fluxes between the cathode channel and gas diffusion layer to be known.

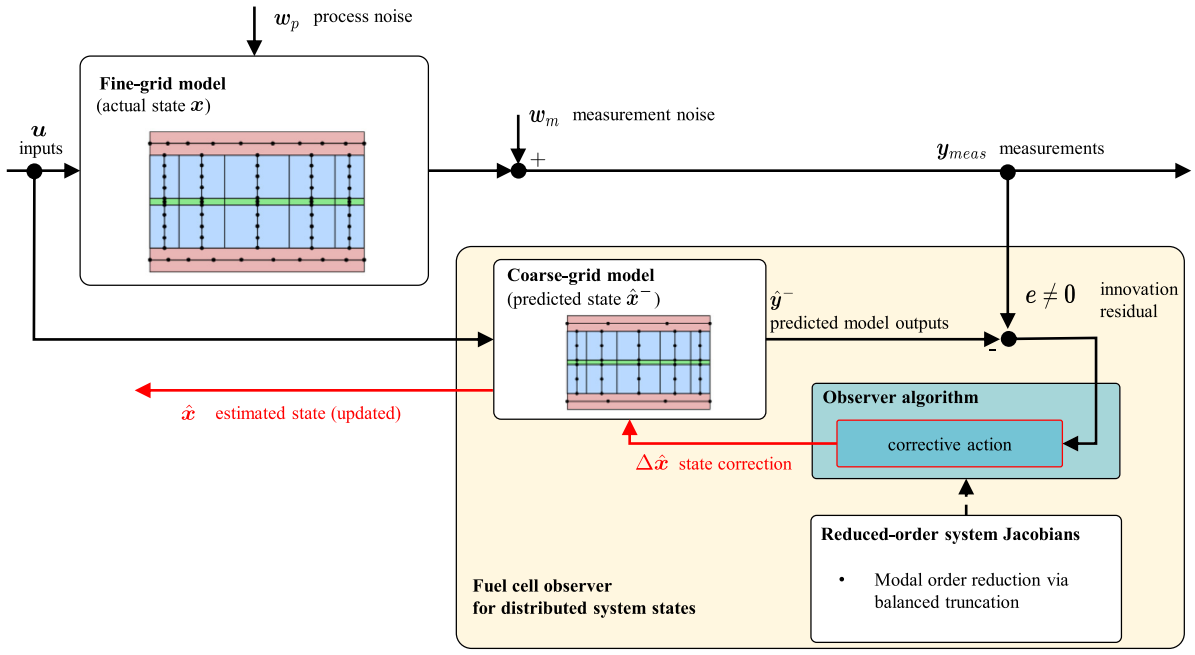


Fig. 1. Scheme of the fuel cell observer for estimating the distributed system state.

The paper is organised as follows. In Section 2 the observation task is explained in more detail, in Section 3 the quasi-2D model is presented, in Section 4 the observer algorithm and model reduction technique are shown, in Section 5 the case study simulation experiment is explained, in Section 6 the results are shown and discussed and the reader is referred to the videos in the supplementary material, and the paper is finally concluded in Section 7.

2. Observation task

The observation task is to estimate the unmeasurable internal states of the system by means of a model, available measurement signals and an appropriate algorithm. The mentioned internal states include: local current density, membrane water content and species concentrations. The model used in this paper is a straight channel isothermal distributed parameter fuel cell model, developed and validated against detailed Computational Fluid Dynamics (CFD) simulations in Murschenhofer et al. (2018). The model overview is given in Section 3. It is important to emphasise that the methodology used in this paper can be applied to practically any type of high-dimensional distributed model, not only to the one used in this work.

The premise is as follows. A fine-grid version of the model is used as simulated reality, against which the observer will be validated. This is necessary as in reality, the validation of an observer can never fully be done as it is not possible to measure the internal states. The fine-grid model provides the measured output signal, in this case, the voltage only. Of course, the input boundary conditions are also known. These include the inlet/outlet pressures, the inlet gas composition, and the current density demand. There is no assumption, however, about the availability of the species concentrations at the system outlet.

Next, a coarse-grid version of the model is used by the observer for the prediction of the internal states. The reason for using a different granularity for the simulated reality and the observer's prediction model is to take into account the plant-model mismatch which is generally always present.

The observer takes in the measured voltage from the simulated reality fine-grid model and at the same time, predicts the voltage using the coarse-grid model. The difference between the two signals, called *innovation residual* is to be minimised. This is done by correcting the internal states of the system. Once that has been achieved, if the system is observable, one can conclude that the internal states computed by the coarse-grid model match reality provided that the model has been validated beforehand so the plant-model mismatch is minimal.

However, a problem arises for a system with many (several hundred or more) states, such as in this case. It is not plausible that all states can freely be corrected for the observer to match the measured voltage. For example, the pressure distribution along the channel should not be in a zig-zag shape just because the observer "thinks" that would be best for minimising the innovation residual. In addition, tuning the observer, i.e. choosing how strongly should each state be available for correction, might become an impossible task for systems with this many states.

It is not plausible that all individual internal states are observable with the available information. For that reason, a model reduction is performed, to detect only the dominant dynamics. Only a few principal *modes* of the system are extracted (see Section 4.2) and the observer makes its correction in the direction of the modes. As a result, the possibly disproportional update of each individual node is avoided. The whole idea described above is visualised in Fig. 1.

In this work, only the voltage is assumed to be a measured output. No additional assumptions of measurements available are made, such as the species concentrations at the system boundaries. The whole procedure results in the estimation of the distribution of internal states of the system.

3. Fuel cell quasi-2D model

This section describes the distributed parameter fuel cell model (Murschenhofer et al., 2018) used as basis for the observer as well as the validation model. The model domain (see Fig. 2) consists of five sections:

1. Cathode gas channel
2. Cathode gas diffusion layer
3. Membrane

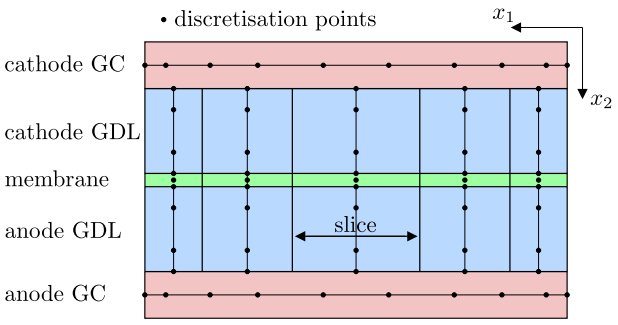


Fig. 2. Spatial discretisation of the quasi-2D fuel cell model.

4. Anode gas diffusion layer
5. Anode gas channel

Every section is coupled via the system equations and appropriate boundary conditions. For modelling details the reader is referred to Murschenhofer et al. (2018). The model is spatially discretised in direction x_1 along-the-channel, and in direction x_2 across the GDL. Therefore, the model is classified as a quasi-2D model.

Some important assumptions are made in the model. First, the model is considered to be isothermal, meaning that the temperature is assumed to be controlled and kept at a constant value of 70 °C. Second, the model does not include liquid water states, but regardless of this fact, membrane dryout/flooding could be avoided if constraints on the membrane water content would be imposed in some control application. Therefore, the information on the membrane water content can be sufficient for practical applications, even if there is no liquid water present in the model.

3.1. Gas channels

Each node in the gas channels is described with the following state variables: gas velocity v, oxygen mass fraction ξ_{O_2} (cathode only), hydrogen mass fraction ξ_{H_2} (anode only), nitrogen mass fraction ξ_{N_2}, water vapour mass fraction ξ_{H_2O}, gas pressure p and gas density ρ. If the number of nodes in the cathode and anode gas channel is N_{cc} and N_{ca}, respectively, then the corresponding state vectors $\boldsymbol{\chi}_{cc}$ and $\boldsymbol{\chi}_{ca}$ are defined as

$$\boldsymbol{\chi}_{cc} = \begin{bmatrix} v_1 \\ \vdots \\ v_{N_{cc}} \\ \xi_{O_2,1} \\ \vdots \\ \xi_{O_2,N_{cc}} \\ \xi_{N_2,1} \\ \vdots \\ \xi_{N_2,N_{cc}} \\ \xi_{H_2O,1} \\ \vdots \\ \xi_{H_2O,N_{cc}} \\ p_1 \\ \vdots \\ p_{N_{cc}} \\ \rho_1 \\ \vdots \\ \rho_{N_{cc}} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\chi}_{ca} = \begin{bmatrix} v_1 \\ \vdots \\ v_{N_{ca}} \\ \xi_{H_2,1} \\ \vdots \\ \xi_{H_2,N_{ca}} \\ \xi_{N_2,1} \\ \vdots \\ \xi_{N_2,N_{ca}} \\ \xi_{H_2O,1} \\ \vdots \\ \xi_{H_2O,N_{ca}} \\ p_1 \\ \vdots \\ p_{N_{ca}} \\ \rho_1 \\ \vdots \\ \rho_{N_{ca}} \end{bmatrix}. \tag{1}$$

3.2. Gas diffusion layers

In the GDLs, the nodes are described with the same variables as in the gas channels. Every internal slice j is described with its own state vector. If the number of nodes in each slice is N_{gc} for the cathode GDL and N_{ga} for the anode GDL, then the cathode and anode GDL state vectors, $\boldsymbol{\chi}_{gc,j}$ and $\boldsymbol{\chi}_{ga,j}$, respectively, are defined analogously as in Eq. (1).

3.3. Membrane

The membrane is described simply by its membrane water content λ in all nodes for every slice. If the number of nodes in the membrane is N_{mem}, then the state vector $\boldsymbol{\chi}_{mem,j}$ for each membrane slice j is defined as

$$\boldsymbol{\chi}_{mem,j} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{N_{mem}} \end{bmatrix}_j. \tag{2}$$

3.4. Local current density and voltage

In addition to all the system variables mentioned until this point, there is the local current density i_j computed for every slice and the cell voltage V, described as an aggregated state.

3.5. State-space system

The full system state vector $\boldsymbol{\chi}_k$ at timestep k is composed of the individual state vectors of the cathode and anode gas channels, N_{sl} slices and the cell voltage. The state vector of every slice is composed of the state vectors of the GDLs and membrane as well as the local current density. The complete system state vector is then

$$\boldsymbol{\chi}_k = \begin{bmatrix} \boldsymbol{\chi}_{cc} \\ \boldsymbol{\chi}_{ca} \\ \boldsymbol{\chi}_{gc,1} \\ \boldsymbol{\chi}_{mem,1} \\ \boldsymbol{\chi}_{ga,1} \\ i_1 \\ \vdots \\ \boldsymbol{\chi}_{gc,N_{sl}} \\ \boldsymbol{\chi}_{mem,N_{sl}} \\ \boldsymbol{\chi}_{ga,N_{sl}} \\ i_{N_{sl}} \\ V \end{bmatrix}_k. \tag{3}$$

The model equations are linearised-in-time (LIT) at every time step.

The LIT technique is explained in detail in Murschenhofer et al. (2018), but for practical reasons, the keypoints of the method are outlined here. The process of discretising the underlying partial differential equations results in a large system of non-linear equations required to be solved each time step. To be computationally efficient, one cannot afford to treat the non-linear terms with numerically expensive iterations. For that reason, the system is successively linearised with respect to the previous time step and only one iteration is necessary to solve the system equations. As a result, the system evolution, i.e. the state vector at time step k, is computed as the solution of a linear system of equations

$$\mathcal{A}(\boldsymbol{\chi}_{k-1}, \boldsymbol{\chi}_{k-2}, \boldsymbol{\chi}_{k-3}, \boldsymbol{u}_{k-1})\boldsymbol{\chi}_k = \mathfrak{b}(\boldsymbol{\chi}_{k-1}, \boldsymbol{\chi}_{k-2}, \boldsymbol{\chi}_{k-3}, \boldsymbol{u}_{k-1}), \tag{4}$$

where $\boldsymbol{u}_{k-1}$ is the system input defined as

$$\boldsymbol{u}_{k-1} = \begin{bmatrix} i_{avg} \\ p_{in}^{cc} \\ p_{out}^{cc} \\ p_{in}^{ca} \\ p_{out}^{ca} \\ \xi_{O_2,in}^{cc} \\ \xi_{N_2,in}^{cc} \\ \xi_{H_2O,in}^{cc} \\ \xi_{H_2,in}^{ca} \\ \xi_{N_2,in}^{ca} \\ \xi_{H_2O,in}^{ca} \end{bmatrix}, \tag{5}$$

consisting of all the boundary conditions mentioned in Section 2. However, only the average current density is dynamically changed and all the other inputs are kept constant.

The sparsity pattern of matrix $\boldsymbol{\mathcal{A}}$ is seen in Fig. 3. The dependence of the matrix $\boldsymbol{\mathcal{A}}$ and vector $\mathfrak{b}$ on the state vector at times $k-1$, $k-2$ and $k-3$ comes from the necessity of computing the time derivatives for the LIT approach. For brevity, the matrices in Eq. (4) will be denoted as $\boldsymbol{\mathcal{A}}_k$ and $\mathfrak{b}_k$. Taking into account the dependence of the system evolution on not only the current state, but also on the last and second last ones, all three time instances ($k-1$, $k-2$ and $k-3$) must be part of the finally complete state vector $\boldsymbol{x}_{k-1} = [\boldsymbol{\chi}_{k-1}^T \ \boldsymbol{\chi}_{k-2}^T \ \boldsymbol{\chi}_{k-3}^T]^T$. Thus, one obtains the state-space formulation of the system

$$\boldsymbol{x}_k = \begin{bmatrix} \boldsymbol{\chi}_k \\ \boldsymbol{\chi}_{k-1} \\ \boldsymbol{\chi}_{k-2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{A}}_k^{-1}\mathfrak{b}_k \\ \boldsymbol{\chi}_{k-1} \\ \boldsymbol{\chi}_{k-2} \end{bmatrix}. \tag{6}$$

By defining

$$\boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}) \equiv \begin{bmatrix} \boldsymbol{\mathcal{A}}_k^{-1}\mathfrak{b}_k \\ \boldsymbol{\chi}_{k-1} \\ \boldsymbol{\chi}_{k-2} \end{bmatrix}, \tag{7}$$

Eq. (6) can be written as

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}) \text{ with } \boldsymbol{x}_k \in \mathbb{R}^{n_x \times 1}. \tag{8}$$

The output equation is

$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k \text{ with } \boldsymbol{y}_k \in \mathbb{R}^{n_y \times 1}. \tag{9}$$

The entries in the output vector $\boldsymbol{y}_k$ are the modelled sensor signals that can be measured and depend on the setup. In this particular case, $\boldsymbol{y}_k = V$, the cell voltage. More detail about the outputs used in this paper are presented in Section 6. Eqs. (8) and (9) describe a nonlinear discrete time state-space model.

3.6. Mole fraction calculation

In the fuel cell community, instead of using the mass fractions ξ, it is more common to represent the species concentrations in mole fractions γ. The oxygen mole fraction is

$$\gamma_{O_2} = \frac{1}{M_{O_2}} \frac{\xi_{O_2}}{\frac{\xi_{O_2}}{M_{O_2}} + \frac{\xi_{N_2}}{M_{N_2}} + \frac{\xi_{H_2O}}{M_{H_2O}}}, \tag{10}$$

where M is the molar mass of each gas. The expression for calculation of the mole fraction for gases other than oxygen is analogous to Eq. (10).

4. Methods

In this section, the observer design as well as the methods for obtaining the reduced-order system Jacobians are presented. All the equations are written in general form and are then applied to the model described in the previous section. This shows the modularity of the algorithm, as it can be used with any high-order nonlinear state-space model. This reduced-dimensionality observer algorithm is visualised in Fig. 4.

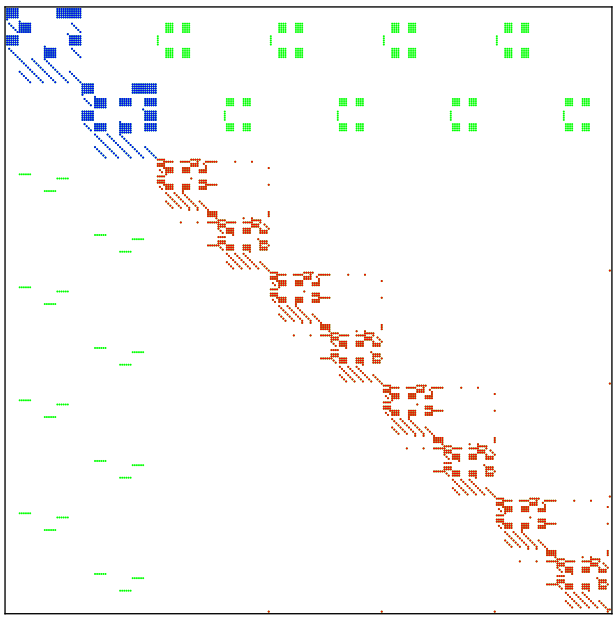


Fig. 3. Sparsity pattern of matrix $\boldsymbol{\mathcal{A}}$ showing the interdependency of the model domains. Every block on the diagonal represents the coupling of the states of one domain. Blue: cathode and anode gas channels; Red: internal slices, every slice is constituted of two GDLs and the membrane in between. The off-diagonal green entries represent the coupling between the slices and gas channels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1. Observer design

An efficient observer design is required to estimate the distributed internal fuel cell states in a robust fashion and in real time, see Fig. 1. In the observer design, the high-order quasi-2D fuel cell model (Section 3) predicts/simulates the distributed system states one time step ahead. Measurements are taken from the simulated reality and compared to the predicted model outputs. If a residual arises, the states of the quasi-2D fuel cell model are updated by an observer algorithm, in this work the proposed reduced-order EKF. The update is computed on the basis of system information given by the sensitivities of the states, the so-called Jacobians. To enhance efficiency, robustness, and convergence of the fuel cell observer algorithm, only the dominant system behaviour is to be considered in the update step.

To do so, the system is transformed from a state-space representation to a modal space (modes $\tilde{\boldsymbol{x}}$), and only the high-energy modes are retained, referred to as $\tilde{\boldsymbol{x}}_r$. The reduced-order system Jacobians that capture the main behaviour in modal space will be denoted as $\tilde{\boldsymbol{A}}_r$. The matrix $\tilde{\boldsymbol{C}}_r$ maps the dominant modes $\tilde{\boldsymbol{x}}_r$ to the model outputs.

In the following section, the formulation of the EKF from Simon (2006) is briefly recapitulated and applied to the present fuel cell problem. A special focus lies on the computation of the update step in modal space.

4.1.1. Background

The Kalman filter is an optimal estimator for linear systems with additive white Gaussian noise and minimises the mean

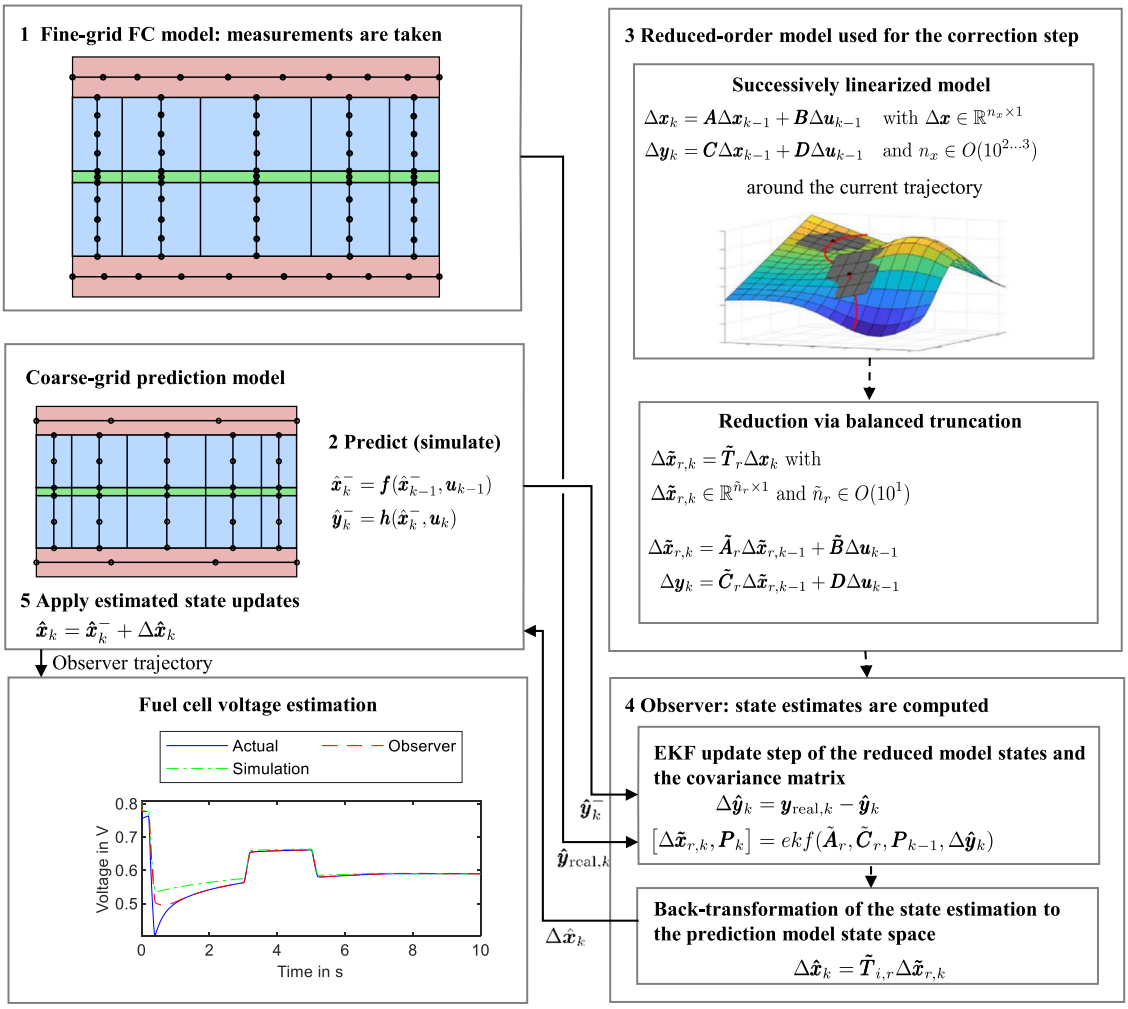


Fig. 4. Reduced-dimensionality observer algorithm.

square errors of the parameters to be estimated. The EKF is one nonlinear extension of it, but the optimality is no longer guaranteed. The discrete-time nonlinear model of the fuel cell (8)–(9) with process noise $\boldsymbol{w}_{\mathrm{p},k}$ and measurement noise $\boldsymbol{w}_{\mathrm{m},k}$ has the form

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1}) + \boldsymbol{w}_{\mathrm{p},k}\,, \tag{11}$$

$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{w}_{\mathrm{m},k}\,. \tag{12}$$

It is assumed that the zero-mean, Gaussian uncorrelated and white noises,

$$\boldsymbol{w}_{\mathrm{p},k} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}) \text{ and } \boldsymbol{w}_{\mathrm{m},k} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})\,, \tag{13}$$

have the known covariance matrices $\boldsymbol{Q}$ and $\boldsymbol{R}$, respectively.

The EKF for estimating the true system states consists of two steps, the prediction step and the update step. The idea of the EKF implementation in this work is to compute the prediction step using the coarse-grid model, and the update step in the modal space using the reduced order Jacobians ($\tilde{\boldsymbol{A}}_{\mathrm{r}}$ and $\tilde{\boldsymbol{C}}_{\mathrm{r}}$).

4.1.2. Prediction step

In the prediction step of the EKF, the predicted state vector, output vector and the error covariance matrix are determined a priori. The predicted state vector $\hat{\boldsymbol{x}}_k^-$ and the predicted output vector $\hat{\boldsymbol{y}}_k^-$ result from evolving the coarse-grid model (8)–(9). The predicted error covariance matrix

$$\tilde{\boldsymbol{P}}_k^- = \tilde{\boldsymbol{A}}_{\mathrm{r}} \tilde{\boldsymbol{P}}_{k-1} \tilde{\boldsymbol{A}}_{\mathrm{r}}^T + \tilde{\boldsymbol{Q}} \tag{14}$$

is obtained in the modal space with the corresponding covariance matrix $\tilde{\boldsymbol{Q}}$. The difference between the predicted output and the measurements $\boldsymbol{y}_{\mathrm{meas}}$ yields the innovation residual,

$$\boldsymbol{e}_k = \boldsymbol{y}_{\mathrm{meas},k} - \hat{\boldsymbol{y}}_k^-\,, \tag{15}$$

and initiates the update step.

4.1.3. Update step

The update step of the EKF is carried out in the reduced modal space. The estimated (corrected) state vector and error covariance matrix are computed a posteriori based on the innovation residual.

First, the innovation covariance and the Kalman correction gain are determined,

$$\tilde{\boldsymbol{S}}_k = \tilde{\boldsymbol{C}}_{\mathrm{r}} \tilde{\boldsymbol{P}}_k^- \tilde{\boldsymbol{C}}_{\mathrm{r}}^T + \boldsymbol{R}\,, \tag{16}$$

$$\tilde{\boldsymbol{K}}_k = \tilde{\boldsymbol{P}}_k^- \tilde{\boldsymbol{C}}_{\mathrm{r}}^T \tilde{\boldsymbol{S}}_k^{-1}\,, \tag{17}$$

respectively. Then, the state correction of the fuel cell model

$$\Delta \hat{\boldsymbol{x}}_k = \mathcal{T} \tilde{\boldsymbol{K}}_k \boldsymbol{e}_k \tag{18}$$

is computed based on the corrective action given by $\tilde{\boldsymbol{K}}_k \boldsymbol{e}_k$ and the transformation from the reduced modal space to the original state-space of the fuel cell model $\mathcal{T}$.

The obtained state correction is applied to the states predicted by the quasi-2D fuel cell model, resulting in the updated

estimated state vector

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^- + \Delta\hat{\boldsymbol{x}}_k \,. \tag{19}$$

The estimation error covariance matrix is updated for the next time step,

$$\tilde{\boldsymbol{P}}_k = \left(\boldsymbol{I} - \tilde{\boldsymbol{K}}_k\tilde{\boldsymbol{C}}_\mathrm{r}\right)\tilde{\boldsymbol{P}}_k^- \,, \tag{20}$$

wherein $\boldsymbol{I}$ is the identity matrix of appropriate size.

The observer's corrective measures are based on the reduced-order system Jacobians (dominant modes only). The high accuracy of the estimated states is achieved through the interaction between the forward simulation of the quasi-2D fuel cell model and the careful measurement-based correction by the observer.

4.2. Reduced-order system Jacobians

The reduced-order system Jacobians capture the main behaviour of the quasi-2D fuel cell model and can be obtained in different ways, one of them being the balanced truncation, which is presented and used in this paper.

4.2.1. Balanced truncation: Background

The balanced truncation is an approach in which the least controllable and observable modes of the system are removed (truncated) and only the dominant behaviour is kept (Moore, 1981). First, the system represented with Eq. (8) is successively linearised around the current point $(\boldsymbol{x}_\mathrm{wp}, u_\mathrm{wp})$. Then, the linearised system is transformed to a balanced realisation, where the controllability and observability Gramians are equal and diagonal (Besselink et al., 2013). Since the diagonal entries of the Gramian reflect the controllability and obervability of the modes, those modes with a small Gramian entry can be truncated. Finally, the truncation step is performed, yielding a reduced order model with the most observable modes only.

4.2.2. Successive linearisation and system scaling

The Extended Kalman Filter approach uses the Jacobians of the nonlinear system defined in Eq. (8). To have a valid correction algorithm in place at all times, the system is successively linearised along the state and input trajectory. The successively linearised model is of the form

$$\boldsymbol{x}_k = \boldsymbol{A}_\mathrm{wp}\boldsymbol{x}_{k-1} + \boldsymbol{B}_\mathrm{wp}u_{k-1} + \boldsymbol{K}_{x,\mathrm{wp}}, \tag{21}$$

where

$$\boldsymbol{A}_\mathrm{wp} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_\mathrm{wp}, \boldsymbol{u}_\mathrm{wp}}, \tag{22}$$

$$\boldsymbol{B}_\mathrm{wp} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{u}_{k-1})}{\partial \boldsymbol{u}}\right|_{\boldsymbol{x}_\mathrm{wp}, \boldsymbol{u}_\mathrm{wp}}, \tag{23}$$

$$\boldsymbol{K}_{x,\mathrm{wp}} = \boldsymbol{f}(\boldsymbol{x}_\mathrm{wp}, \boldsymbol{u}_\mathrm{wp}) - \boldsymbol{A}_\mathrm{wp}\boldsymbol{x}_\mathrm{wp} - \boldsymbol{B}_\mathrm{wp}\boldsymbol{u}_\mathrm{wp} \,. \tag{24}$$

In order to compute the Gramians and transform the system to its balanced realisation, it is necessary to prescale the system. The scaled state vector $\boldsymbol{x}_\mathrm{s}$ is then defined as $\boldsymbol{x}_\mathrm{s} = \boldsymbol{X}_\mathrm{s}^{-1}\boldsymbol{x}$, where $\boldsymbol{X}_\mathrm{s}$ is a diagonal square matrix. The same procedure is applied to the input and output vectors: $u_\mathrm{s} = U_\mathrm{s}^{-1}u$ and $\boldsymbol{y}_\mathrm{s} = \boldsymbol{Y}_\mathrm{s}^{-1}\boldsymbol{y}$, respectively.

The index "s" stands for "scaled".

4.2.3. Balanced realisation

Using the scaled system, the balanced realisation can be computed. To cope with the off-equilibrium term $K_{x,\mathrm{wp,s}}$, it is lumped together with the input matrix in the following way,

$$\boldsymbol{x}_{\mathrm{s},k} = \boldsymbol{A}_{\mathrm{s,wp}}\boldsymbol{x}_{\mathrm{s},k-1} + \begin{bmatrix}\boldsymbol{B}_{\mathrm{s,wp}} & \boldsymbol{K}_{x,\mathrm{wp,s}}\end{bmatrix}\begin{bmatrix}u_{\mathrm{s},k-1}\\ 1\end{bmatrix}, \tag{25}$$

formulating the system in a suitable way for the MATLAB command `balreal`, used to compute the balanced realisation.

The state transformation

$$\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{T}}\boldsymbol{x}_\mathrm{s} \tag{26}$$

is used to convert system (25) into its balanced realisation, in which the controllability and observability Gramians are equal and diagonal. In Eq. (26), $\tilde{\boldsymbol{x}} \in \mathbb{R}^{n_x\times 1}$ and $\tilde{\boldsymbol{T}} \in \mathbb{R}^{n_x\times n_x}$.

The balanced state vector $\tilde{\boldsymbol{x}}$ is partitioned into two parts: the one that is going to be retained $\tilde{\boldsymbol{x}}_\mathrm{r}$ for the reduced-order system and the other, which is going to be eliminated $\tilde{\boldsymbol{x}}_\mathrm{e}$,

$$\tilde{\boldsymbol{x}} = \begin{bmatrix}\tilde{\boldsymbol{x}}_\mathrm{r}\\ \tilde{\boldsymbol{x}}_\mathrm{e}\end{bmatrix}, \tag{27}$$

where $\tilde{\boldsymbol{x}}_\mathrm{r} \in \mathbb{R}^{\tilde{n}_\mathrm{r}\times 1}$ and $\tilde{\boldsymbol{x}}_\mathrm{e} \in \mathbb{R}^{\tilde{n}_\mathrm{e}\times 1}$, with $\tilde{n}_\mathrm{r} + \tilde{n}_\mathrm{e} = n_x$, the total number of states. The truncation value $\tilde{n}_\mathrm{r}$ is chosen upon analysing the Gramian of the system. Following this notation, the state transformation matrix $\tilde{\boldsymbol{T}}$ and its inverse $\tilde{\boldsymbol{T}}_\mathrm{i} = \tilde{\boldsymbol{T}}^{-1}$ are also partitioned into the part which maps the true states to the reduced order ones and the part which maps them to the eliminated ones. Inserting Eq. (27) into (26) and performing the partition of the transformation matrices, the direct mapping between the reduced and eliminated states, and the true states, is obtained,

$$\begin{bmatrix}\tilde{\boldsymbol{x}}_\mathrm{r}\\ \tilde{\boldsymbol{x}}_\mathrm{e}\end{bmatrix} = \begin{bmatrix}\tilde{\boldsymbol{T}}_\mathrm{r}\\ \tilde{\boldsymbol{T}}_\mathrm{e}\end{bmatrix}\boldsymbol{x}_\mathrm{s}, \tag{28}$$

as well as its inverse

$$\boldsymbol{x}_\mathrm{s} = \begin{bmatrix}\tilde{\boldsymbol{T}}_{\mathrm{i,r}} & \tilde{\boldsymbol{T}}_{\mathrm{i,e}}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{x}}_\mathrm{r}\\ \tilde{\boldsymbol{x}}_\mathrm{e}\end{bmatrix}, \tag{29}$$

where $\tilde{\boldsymbol{T}}_\mathrm{r} \in \mathbb{R}^{\tilde{n}_\mathrm{r}\times n_x}$, $\tilde{\boldsymbol{T}}_\mathrm{e} \in \mathbb{R}^{\tilde{n}_\mathrm{e}\times n_x}$, $\tilde{\boldsymbol{T}}_{\mathrm{i,r}} \in \mathbb{R}^{n_x\times \tilde{n}_\mathrm{r}}$ and $\tilde{\boldsymbol{T}}_{\mathrm{i,e}} \in \mathbb{R}^{n_x\times \tilde{n}_\mathrm{e}}$.

4.2.4. Truncation

Since they contribute much less to the system behaviour, the states $\tilde{\boldsymbol{x}}_\mathrm{e}$ are simply deleted and the system is described with $\tilde{\boldsymbol{x}}_\mathrm{r}$ only, yielding the following transformation for the reduced states,

$$\tilde{\boldsymbol{x}}_\mathrm{r} = \tilde{\boldsymbol{T}}_\mathrm{r}\boldsymbol{x}_\mathrm{s}. \tag{30}$$

Inverting Eq. (30) and inserting it into Eq. (25) and then left multiplying the whole expression with $\tilde{\boldsymbol{T}}_\mathrm{r}$, one obtains the reduced-order system

$$\tilde{\boldsymbol{x}}_{\mathrm{r},\,k} = \underbrace{\tilde{\boldsymbol{T}}_\mathrm{r}\boldsymbol{A}_{\mathrm{s,wp}}\tilde{\boldsymbol{T}}_{\mathrm{i,r}}}_{\tilde{\boldsymbol{A}}_\mathrm{r}}\tilde{\boldsymbol{x}}_{\mathrm{r},\,k-1} + \tilde{\boldsymbol{T}}_\mathrm{r}\begin{bmatrix}\boldsymbol{B}_{\mathrm{s,\,wp}} & \boldsymbol{K}_{x,\mathrm{wp,s}}\end{bmatrix}\begin{bmatrix}u_{\mathrm{s},k-1}\\ 1\end{bmatrix}. \tag{31}$$

Finally, the Jacobian $\tilde{\boldsymbol{A}}_\mathrm{r} = \tilde{\boldsymbol{T}}_\mathrm{r}\boldsymbol{A}_{\mathrm{s,wp}}\tilde{\boldsymbol{T}}_{\mathrm{i,r}}$ is obtained, with $\tilde{\boldsymbol{A}}_\mathrm{r} \in \mathbb{R}^{\tilde{n}_\mathrm{r}\times\tilde{n}_\mathrm{r}}$. Also, the new output equation is

$$\boldsymbol{y}_{\mathrm{s},k} = \tilde{\boldsymbol{C}}_\mathrm{r}\tilde{\boldsymbol{x}}_{\mathrm{r},k} \text{ with } \tilde{\boldsymbol{C}}_\mathrm{r} \in \mathbb{R}^{n_y\times\tilde{n}_\mathrm{r}} \,, \tag{32}$$

where $\tilde{\boldsymbol{C}}_\mathrm{r} = \boldsymbol{C}_{\mathrm{s,wp}}\tilde{\boldsymbol{T}}_{\mathrm{i,r}}$.

The Jacobian $\tilde{\boldsymbol{A}}_\mathrm{r}$ is used in the update step of the EKF to correct the dominant modes. To obtain the correction in the original states, the modes are projected onto the states using the transformation matrix $\mathcal{T}$ from Eq. (18). For the approach described above, $\mathcal{T}$ is defined as:

$$\mathcal{T} = \boldsymbol{X}_\mathrm{s}\tilde{\boldsymbol{T}}_{\mathrm{i,r}}. \tag{33}$$

4.3. Observability analysis

The observability of the system is approached from two different angles. One is the observability of the modes and the other is the observability of the states. In Table 4, the different model sizes

Table 4
Different model sizes.

Model	N_{cc}	N_{gc}	N_{mem}	N_{ga}	N_{ca}	N_{sl}	Δt/ms
Fine	30	8	5	8	30	10	2
Coarse	6	4	5	4	6	4	6.25

of the fine-grid and coarse-grid model are shown. Using this data and the state vector structure from Eqs. (1)–(3), one can compute the total number of states n_x in each model:

$$n_x = 3\left(6N_{cc} + 6N_{sl}N_{gc} + N_{sl}N_{mem} + N_{sl} + 6N_{sl}N_{ga} + 6N_{ca} + 1\right). \tag{34}$$

The factor 3 comes from the fact that three time instances, $k-1$, $k-2$ and $k-3$ are included in the state vector for the LIT approach, as explained in Section 3. The number of states of the fine-grid model is 4143 and for the coarse-grid model the number of states is 867. However, observability is analysed only for the first third of the state vector and only for the coarse-grid model since the fine-grid model serves the role of simulated reality. Therefore, there are 289 states in the first third of the state vector of the coarse-grid model. The number of modes $\tilde{n}_\mathrm{r}$ obtained by balanced truncation is chosen to be 5.

4.3.1. Observability of the modes

The observability of the modes is computed by checking the rank of the observability matrix Q_{ob} at any operating point defined as:

$$Q_{ob} = \begin{bmatrix} \tilde{\boldsymbol{C}}_\mathrm{r} \\ \tilde{\boldsymbol{C}}_\mathrm{r}\tilde{\boldsymbol{A}}_\mathrm{r} \\ \tilde{\boldsymbol{C}}_\mathrm{r}\tilde{\boldsymbol{A}}_\mathrm{r}^2 \\ \vdots \\ \tilde{\boldsymbol{C}}_\mathrm{r}\tilde{\boldsymbol{A}}_\mathrm{r}^{\tilde{n}_\mathrm{r}-1} \end{bmatrix} \tag{35}$$

The observability matrix Q_{ob} is shown to have full rank for any tested operating point meaning that the principal modes of the system are observable.

4.3.2. Observability of the states

The observability of the states cannot be easily computed via the observability matrix as in the case for the modes due to the large number of states. One would have to compute high powers of the Jacobian matrix $\boldsymbol{A}_\mathrm{wp}$ which becomes numerically unreliable. Instead, the observability of the states is analysed by using the transformation matrix $\tilde{\boldsymbol{T}}_\mathrm{r}$. This matrix maps the states to the modes and by analysing the rows of this matrix, one can see in which direction in the state space all the dominant modes point to. In Fig. 5, one can see the entries of the rows of this matrix from top to bottom, starting with the most dominant mode on the top. Any location on the x-axis corresponds to a certain state and the value on the y-axis is the absolute value of the entry of each row of the matrix $\tilde{\boldsymbol{T}}_\mathrm{r}$. Upon analysing the rows of this matrix, one can see that the modes of the system point in the direction of all domains, but only for the anode channel it is very weak, indicating that all domains except the anode channel are observable. The modes do not point in the direction of all states, but there is neither the need for it, as per the strong coupling between the adjacent states, the profiles will follow the observer's corrective action. To remedy the anode channel observability problem, it must be pointed out that the analysis of the transformation matrix entries must be supplemented with information about the known boundary conditions. Namely, for the anode channel, the inlet gas composition and pressure as well as the outlet pressure are known which gives enough information for the system to be fully observable, as confirmed by the results in Section 6. To elaborate further, these values at the boundary are also part of the state vector, but the time evolution of these states is not computed, but imposed as an input, and that is the reason why in the transformation matrix entries, they do not show up since the observer will not correct them, but they are, of course, known, providing good observability.

5. Simulation experiment

In this section, the simulation scenario of the case study mentioned in the introduction is explained in detail. The performance of the observer is demonstrated using a more detailed simulation, considered as reality, to evaluate the convergence of all states inside the fuel cell. An evaluation of the observer based on real measurements is currently impossible, as no sensors are available to detect the distributed state inside the fuel cell.

5.1. Setup

The goal of the observer is to estimate all states of the fuel cell given in Eq. (3). In this work, special attention is paid to the along-the-channel spatial distribution of the following quantities:

- Current density i
- Membrane water content λ
- Species mole fractions in the cathode and anode gas channels γ_j, where $j \in \{O_2, N_2, H_2O\}$ for the cathode and $j \in \{H_2, N_2, H_2O\}$ for the anode

It is assumed that only a few inexpensive sensors are available to the observer for measurements, namely cell voltage and pressure at the system boundaries. In this experiment, the system is considered pressure-driven, therefore, the pressures at the system boundaries are set as boundary conditions, as seen in Eq. (5). The output of the system, available to the observer as measurement is the cell voltage $\boldsymbol{y}_\mathrm{meas} = V$. If the fuel cell would be operated in a mass-driven mode, then the pressures at the system boundaries would also constitute the output vector. The measurements are subject to Gaussian noise with a standard deviation of 1% of the corresponding quantity. The system is excited with a time-varying input, the average current density, whose signal is seen in Fig. 6.

5.2. Simulation models and their resolution

Three simulation runs are performed:

- Measurement generation using a fine-grid model, referred to as *Actual*
- Reference simulation using a coarse-grid model without the observer, referred to as *Simulation*
- Estimation using a coarse-grid model with observer, referred to as *Observer*

The number of discretisation points, illustrated in Fig. 2, determine the resolution of the model. In the following simulation study, models of different sizes are used. First, a fine-grid model is used to generate the simulated reality that the observer will try to estimate, i.e. the *Actual* states distribution. The information on the internal states of this first simulation run is by no means known to the observer, which receives only information about the system outputs. The model used by the observer is a coarse-grid model, meaning that it has less discretisation points than the reality model, introducing a model error. The models' grid sizes and the time step size Δt are given in Table 4. To show the performance of the observer, it is always interesting to see what

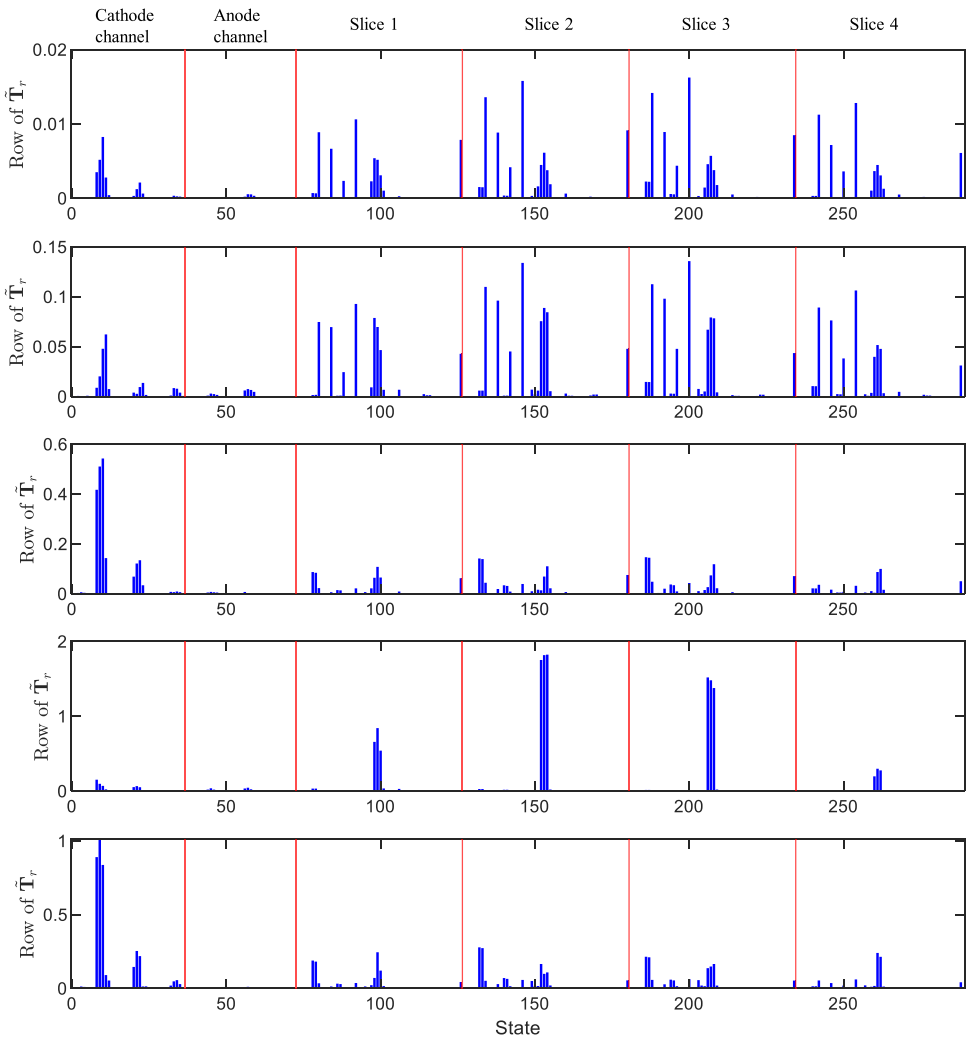


Fig. 5. Entries of the rows of the transformation matrix $\tilde{\boldsymbol{T}}_{\mathrm{r}}$. The top plot shows the entries for the first dominant mode, and so on until the bottom. The entries show the direction in which the observer corrects the system. The red lines separate the domains in the state vector. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

would happen in a pure simulation, in which the observer is not active. In this *Simulation* run, the system is initialised differently than in the first run, as it is not possible to know the initial state conditions in reality. Also, the observer is not active. If the system is asymptotically stable, as it is the case for this model, the system will eventually converge towards the true state, regardless of the wrong initialisation. However, if the observer is active, it is aware of the error between simulation and reality and tries to match the states accordingly. In this third *Observer* simulation run with wrong initialisation, the system converges towards the true state faster, which is crucial for diagnostics and/or fault detection, especially if the initial error is large or if slow dynamics are present in the system. In Fig. 7, the voltage behaviour in all three simulation runs is shown as demonstration of the explained procedure.

6. Results and discussion

The results of the simulation are presented both in the form of figures in this paper as well as two videos which can be found in the supplementary material. The system is excited with the average current density demand as shown in Fig. 6. The vertical magenta lines in Figs. 6 and 7 mark the times for which the snapshots, i.e., Figs. 8 to 10 are plotted. In these figures, the top

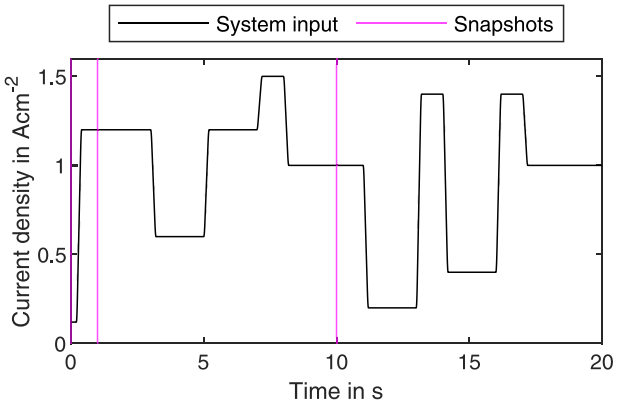


Fig. 6. Excitation signal. The solid black line is the input (average current density) trajectory and the magenta lines indicate the times where snapshots shown in Section 6 are taken. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

plot is the system snapshot at the initial time $t = 0$ s, the middle plot shows the system at time $t = 1$ s and the bottom plot shows the system at time $t = 6$ s. The purpose of the figures is to give

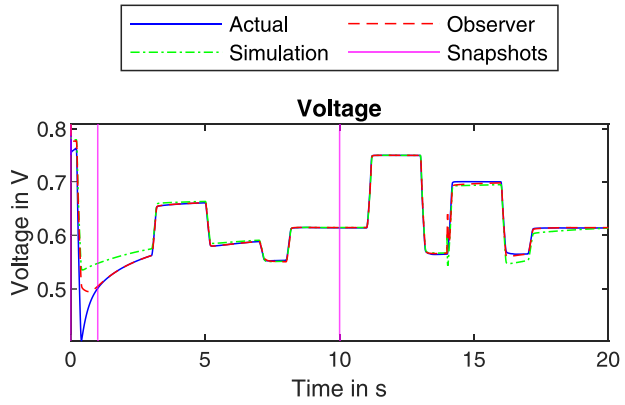


Fig. 7. Cell voltage. The observer (red line) converges towards the true value (blue line) faster than the uncorrected simulation (green line). The magenta lines indicate the times where snapshots shown in Section 6 are taken. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

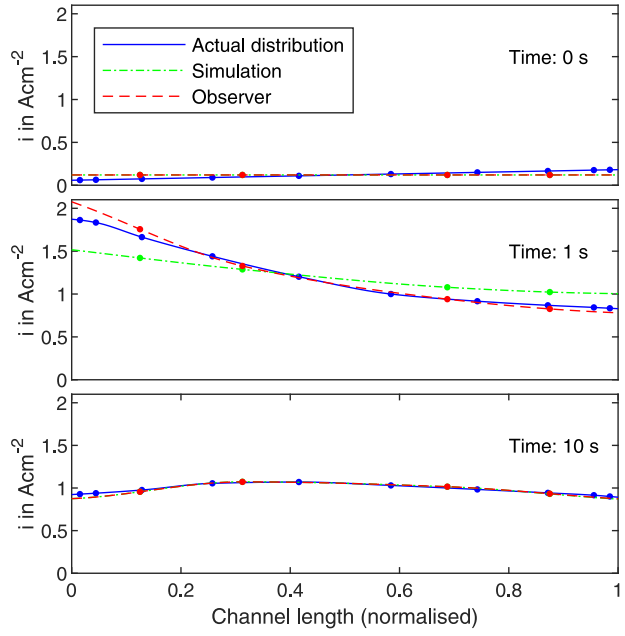


Fig. 8. Local current density distribution along the channel at different times. The reality model has 10 and the simulation model has 4 nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the opportunity to visualise the procedure at a glance and to point out some interesting details, but the whole richness of the results is found in Videos 1 and 2.

6.1. Local current density and membrane water content

The current density distribution along the channel is shown in Video 1 (top plot) as well as in Fig. 8. The number of plotted nodes for the simulated reality fine-grid model is 10, equal to the number of slices (blue line). The simulation without the observer is plotted in green and the simulation where the observer is active is plotted in red. The number of nodes in the green and red lines is 4, equal to the number of slices of the coarse-grid model. The profile of all the states is interpolated with a Piecewise Cubic Hermite Interpolating Polynomial using the MATLAB `pchip` command. The current density distribution in the simulation model is

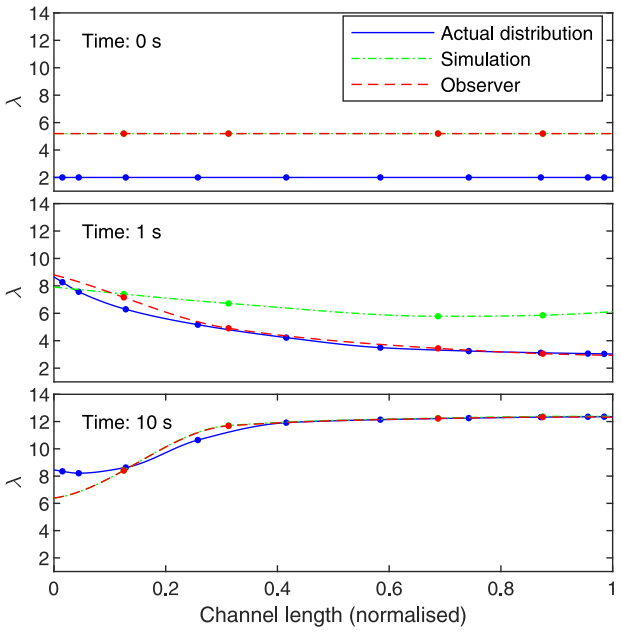
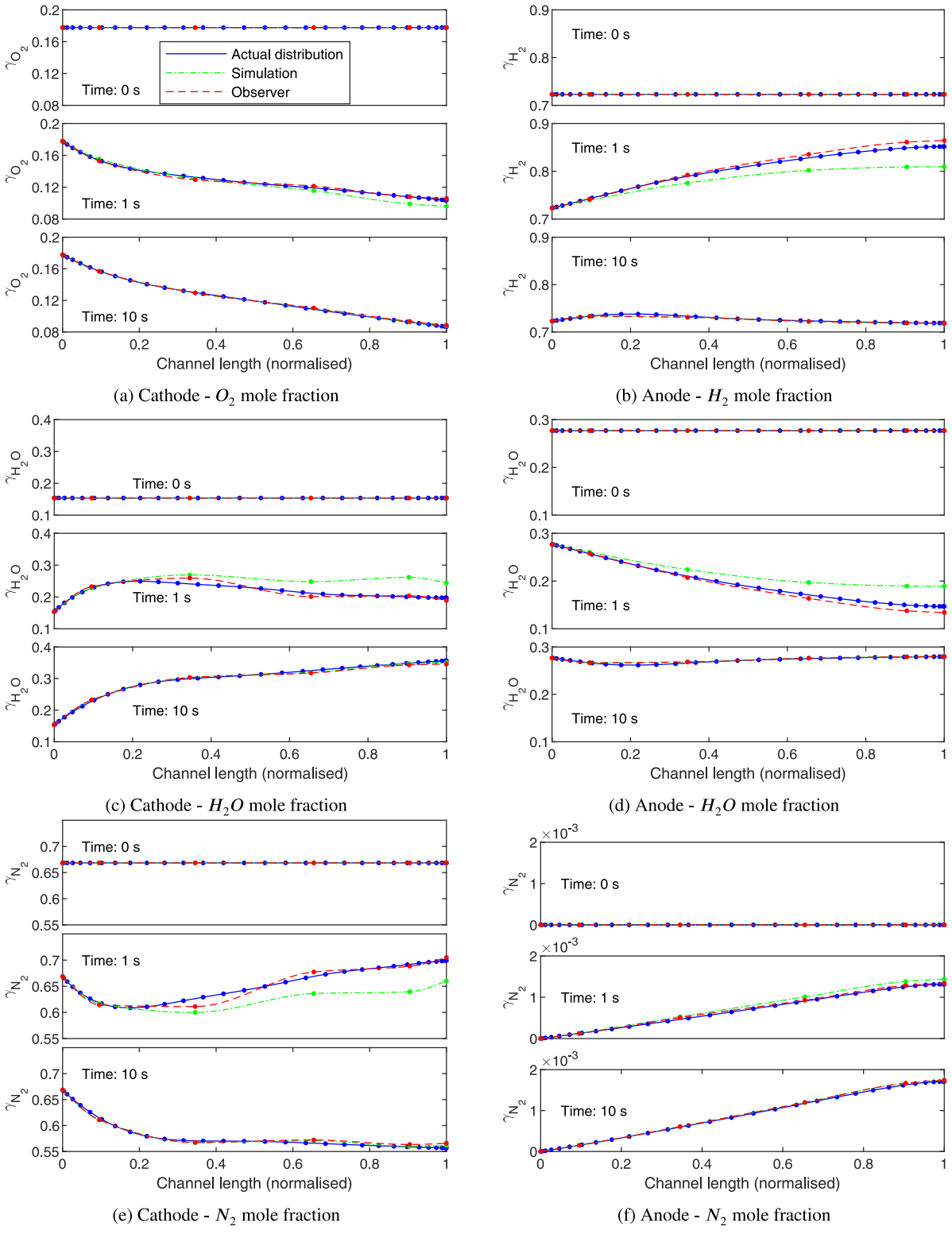


Fig. 9. Membrane water content distribution along the channel at different times. The observer converges towards the true state much faster than the uncorrected simulation.

initialised differently from reality, as it is not realistic to assume the initial distribution in reality to be known. In Video 1, the initialisation can be seen at time $t = 0$ s as well as in the top plot of Fig. 8. At time $t = 1$ s, in the middle plot of Fig. 8, one can see that the local current density distribution estimated by the observer already almost matches the true state. Furthermore, it is to be noted that the whole profile of the local current density distribution matches reality even though the coarse-grid model has fewer nodes and the nodes position is not equal to the ones in the fine-grid model. This highlights the robustness of the algorithm. At time $t = 6$ s, seen in the bottom plot of Fig. 8 both the simulation and the observer have converged towards the true state completely. The membrane water content is shown in Video 1 (middle plot) and Fig. 9. As with the current density, in the simulation, it is initialised differently than reality, as seen in Video 1 at time $t = 0$ s as well as in the top plot of Fig. 9. The main reason for the big voltage offset seen in Fig. 7 at the beginning of the simulation is exactly the membrane water content. The membrane is assumed to be too humid at this point and the observer corrects the error by dehumidifying the membrane quite fast, whereas the pure simulation takes long to dehumidify the membrane due to the slow water transport dynamics. This can be clearly seen in the middle and bottom plot of Fig. 9. In the figures, however, the transition from the initial value to the true state cannot be captured, but in Video 1, the transient can be clearly seen.

6.2. Species mole fractions

The second group of results are the species mole fractions shown in Fig. 10 and in Video 2. Unlike the local current density and membrane water content, the species mole fractions are initialised the same way in the fine-grid and coarse-grid models, but as already stated, the main effect for the voltage deviation is the error in the membrane water content. Since the membrane starts with a significantly higher water content in the simulation and observer than in the fine-grid model, the water is driven towards the channel at different rates because the concentration gradient



(a) Cathode - O_2 mole fraction

(b) Anode - H_2 mole fraction

(c) Cathode - H_2O mole fraction

(d) Anode - H_2O mole fraction

(e) Cathode - N_2 mole fraction

(f) Anode - N_2 mole fraction

Fig. 10. The distribution of the species mole fractions along the cathode (a, c, e) and anode (b, d, f) channels. The actual distributions, denoted in blue, are generated by a high-fidelity fine-grid model. The green lines denote simulations of a coarse-grid model with no feedback from sensor data. The red lines indicate the distribution estimates generated by the proposed observer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

is different. In addition, the difference between simulation and observer comes from the corrective action of the observer. Very quickly, though, the observer corrects the mole fractions, visible in the middle and bottom plots of the state triplets in Fig. 10. There are 30 nodes in the gas channels for the reality model and only 6 for the simulation model. An interesting detail can be observed in this experiment. For example, in the middle plot of Fig. 10(c), the water mole fraction in the cathode is seen. If the mole fraction at the end of the channel could be measured, and no observer existed, one would see that the simulation deviates

at that point by some value, but the actual error is larger than that — a significant portion of the whole profile is wrong. One would be ignorant to this fact from measuring the mole fraction at the outlet only. On the other hand, with an observer in place, one can estimate the distribution better, since by its very nature, the observer pushes the system towards reality.

7. Conclusion

Monitoring the internal states of a fuel cell is challenging but crucial for its safe and lasting operation. While measuring them is almost impossible or only realisable with highly expensive equipment, observers offer a good alternative to estimate the distributed states on the basis of models and a few simple measurements. However, realistic modelling of fuel cells leads to complex high-order systems described by partial differential equations. The implementation of such models in observers is not straightforward when keeping real-time feasibility and robustness in mind. In this work, we have developed an observer that uses a full fuel cell model to predict future states and reduced system Jacobians (dominant modes only) to update the predicted states based on measurements. The method has advantages in terms of good observability and computational efficiency, as only the dominant behaviour is considered in the correction of the predicted states, resulting in well-converging state estimates.

The developed method is tested using a high-fidelity simulated reality to best evaluate the convergence of the distributed system states. A model with wrong initialisation and model errors (coarser discretisation in space and time) is used as a reference simulation and serves as the basis for the observer. Only a few realistic and inexpensive sensors at the inlet and outlet of the fuel cell provide the observer with measurements. It turns out that the observer updates the simulation model correctly and quickly drives all states to their true values. The estimates of the distributed internal states converge well, allowing monitoring of critical states in the fuel cell. A pure simulation without observer correction leads to slow convergence and is not useful to improve the operation of a fuel cell.

The novel fuel cell observer is ready to be implemented in real-world applications, such as the automotive industry, where critical conditions may occur due to transient and dynamic operations and highly expensive laboratory measurement equipment is not available. Currently, the observer algorithm is being extended to include parameter estimation in addition to state estimation. This means that an imperfect model will correct itself, relaxing the need for perfect initial parametrisation, which would require expensive experimental tests. In future work, the observer can be extended with additional functionalities, e.g. estimation of changing parameters due to ageing processes.

CRediT authorship contribution statement

Martin Vrlić: Conceptualization, Methodology, Software, Simulations, Writing – original draft. **Dominik Pernsteiner:** Conceptualization, Methodology, Software, Simulations, Writing – original draft, Visualisation. **Alexander Schirrer:** Writing – reviewing and editing, Supervision. **Christoph Hametner:** Supervision, Funding acquisition. **Stefan Jakubek:** Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was funded by the FFG project AlpeDHues (grant number 884322) and the Christian Doppler Research Association. The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding programme.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.egyr.2023.06.006.

References

Benner, P., Gugercin, S., Willcox, K., 2015. A survey of projection-based model reduction methods for parametric dynamical systems. SIAM Rev. 57, 483–531. http://dx.doi.org/10.1137/130932715.

Besselink, B., Tabak, U., Lutowska, A., Wouw, N.V.D., Nijmeijer, H., Rixen, D.J., Hochstenbach, M.E., Schilders, W.H.A., 2013. A comparison of model reduction techniques from structural dynamics, numerical mathematics and systems and control. J. Sound Vib. 332, 4403–4422. http://dx.doi.org/10.1016/j.jsv.2013.03.025.

Brunton, S.L., Kutz, J.N., 2019. Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control. Cambridge University Press, Cambridge (U.K.), http://dx.doi.org/10.1017/9781108380690.

Dobrokhotov, V., Larin, A., 2019. Multisensory gas chromatography for field analysis of complex gaseous mixtures. ChemEngineering 3, 1–18. http://dx.doi.org/10.3390/chemengineering3010013.

Farrell, B.F., Ioannou, P.J., 2001. State estimation using a reduced-order Kalman filter. J. Atmos. Sci. 58, 3666–3680. http://dx.doi.org/10.1175/1520-0469(2001)058{<}3666:SEUARO{>}2.0.CO;2.

Hidayat, Z., Babuska, R., De Schutter, B., Núñez, A., 2011. Observers for linear distributed-parameter systems: A survey. In: 2011 IEEE International Symposium on Robotic and Sensors Environments. ROSE, pp. 166–171. http://dx.doi.org/10.1109/ROSE.2011.6058523.

IEA, 2019. The future of hydrogen. URL: https://www.iea.org/reports/the-future-of-hydrogen. (Accessed 14 January 2022).

IEA, 2021. Global hydrogen review 2021. URL: https://www.iea.org/reports/global-hydrogen-review-2021. (Accessed 14 January 2022).

İnci, M., Büyük, M., Demir, M.H., İlbey, G., 2021. A review and research on fuel cell electric vehicles: Topologies, power electronic converters, energy management methods, technical challenges, marketing and future aspects. Renew. Sustain. Energy Rev. 137, http://dx.doi.org/10.1016/j.rser.2020.110648.

Jazwinski, A., 1970. Stochastic Processes and Filtering Theory. Academic Press, New York (NY, USA) and London (U.K.).

Khodadadi, H., Jazayeri-Rad, H., 2011. Applying a dual extended Kalman filter for the nonlinear state and parameter estimations of a continuous stirred tank reactor. Comput. Chem. Eng. 35, 2426–2436. http://dx.doi.org/10.1016/j.compchemeng.2010.12.010.

Lee, J., Nam, O., Cho, B.H., 2007. Li-ion battery SOC estimation method based on the reduced order extended Kalman filtering. J. Power Sources 174, 9–15. http://dx.doi.org/10.1016/j.jpowsour.2007.03.072.

Lee, J., Nguyen, H.D., Escribano, S., Micoud, F., Rosini, S., Tengattini, A., Atkins, D., Gebel, G., Iojoiu, C., Lyonnard, S., Morin, A., 2021. Neutron imaging of operando proton exchange membrane fuel cell with novel membrane. J. Power Sources 496, http://dx.doi.org/10.1016/j.jpowsour.2021.229836.

Luna, J., Husar, A., Serra, M., 2015. Nonlinear distributed parameter observer design for fuel cell systems. Int. J. Hydrogen Energy 40, 11322–11332. http://dx.doi.org/10.1016/j.ijhydene.2015.05.132.

Luna, J., Jemei, S., Yousfi-Steiner, N., Husar, A., Serra, M., Hissel, D., 2016a. Nonlinear predictive control for durability enhancement and efficiency improvement in a fuel cell power system. J. Power Sources 328, 250–261. http://dx.doi.org/10.1016/j.jpowsour.2016.08.019.

Luna, J., Usai, E., Husar, A., Serra, M., 2016b. Nonlinear observation in fuel cell systems: A comparison between disturbance estimation and high-order sliding-mode techniques. Int. J. Hydrogen Energy 41, 19737–19748. http://dx.doi.org/10.1016/j.ijhydene.2016.06.041.

Luna, J., Usai, E., Husar, A., Serra, M., 2017. Enhancing the efficiency and lifetime of a proton exchange membrane fuel cell using nonlinear model-predictive control with nonlinear observation. IEEE Trans. Ind. Electron. 64, 6649–6659. http://dx.doi.org/10.1109/TIE.2017.2682787.

Moore, B., 1981. Principal component analysis in linear systems: Controllability, observability, and model reduction. IEEE Trans. Automat. Control 26, 17–32. http://dx.doi.org/10.1109/TAC.1981.1102568.

Murschenhofer, D., Kuzdas, D., Braun, S., Jakubek, S., 2018. A real-time capable quasi-2D proton exchange membrane fuel cell model. Energy Convers. Manage. 162, 159–175. http://dx.doi.org/10.1016/j.enconman.2018.02.028.

Park, B.-G., Kim, J.-M., Kim, J.-W., Lee, K.-C., Koo, D.-H., Hyun, D.-S., 2013. New approach to EKF-based sensorless control using parallel structure for non-salient pole permanent magnet synchronous motors. In: 2013 International Conference on Electrical Machines and Systems. ICEMS, pp. 1099–1104. http://dx.doi.org/10.1109/ICEMS.2013.6754407.

Pernsteiner, D., Schirrer, A., Kasper, L., Hofmann, R., Jakubek, S., 2021. State estimation concept for a nonlinear melting/solidification problem of a latent heat thermal energy storage. Comput. Chem. Eng. 153, 107444. http://dx.doi.org/10.1016/j.compchemeng.2021.107444.

Sarmiento-Carnevali, M., Serra, M., Batlle, C., 2017. Distributed parameter model-based control of water activity and concentration of reactants in a polymer electrolyte membrane fuel cell. Int. J. Hydrogen Energy 42, 26389–26407. http://dx.doi.org/10.1016/j.ijhydene.2017.08.191.

Schmid, P.J., 2010. Dynamic mode decomposition of numerical and experimental data. J. Fluid Mech. 656, 5–28. http://dx.doi.org/10.1017/S0022112010001217.

Simon, D., 2006. Optimal State Estimation: Kalman, H∞, and Nonlinear Approaches. John Wiley & Sons, Hoboken (NJ, USA), http://dx.doi.org/10.1002/0470045345.

Wang, X.R., Ma, Y., Gao, J., Li, T., Jiang, G.Z., Sun, Z.Y., 2021. Review on water management methods for proton exchange membrane fuel cells. Int. J. Hydrogen Energy 46, 12206–12229. http://dx.doi.org/10.1016/j.ijhydene.2020.06.211.

Wishner, R.P., Tabaczynski, J.A., Athans, M., 1969. A comparison of three non-linear filters. Automatica 5, 487–496. http://dx.doi.org/10.1016/0005-1098(69)90110-1.

Yuan, H., Dai, H., Wei, X., Ming, P., 2020. Model-based observers for internal states estimation and control of proton exchange membrane fuel cell system: A review. J. Power Sources 468, http://dx.doi.org/10.1016/j.jpowsour.2020.228376.

Zaghloul, M.A.S., Mason, J.H., Wang, M., Buric, M., Peng, Z., Lee, S., Ohodnicki, P., Abernathy, H., Chen, K.P., 2021. High spatial resolution temperature profile measurements of solid-oxide fuel cells. Appl. Energy 288, http://dx.doi.org/10.1016/j.apenergy.2021.116633.