

# Data-Driven Active Power Flow Modeling: A Behavioral Systems Approach

Sebastian Otzen<sup>1</sup>, Hannes M. H. Wolf<sup>2</sup> and Christian A. Hans<sup>2</sup>

**Abstract**—The increasing decentralization of power systems driven by a large number of renewable energy sources poses challenges in power flow optimization: Partially unknown power line properties can render model-based approaches unsuitable. With the increasing deployment of sensors, data-driven methods rise as a promising alternative, offering flexibility to adapt to changes and deal with unknown properties. In this paper, we propose a novel data-driven representation of nonlinear active power flow equations for radial grids based on Willems’ Fundamental Lemma. Our approach allows for direct integration of input/output data into active power flow optimization, enabling cost minimization and constraint enforcement without requiring explicit knowledge of power line properties or a separate system identification step. Moreover, we derive a convex relaxation and show in a numerical case study that our approaches yield results identical to those of optimal active power flow formulations with known parameters.

## I. INTRODUCTION

Power systems are shifting from centralized generation to decentralized, small-scale units such as photovoltaic generators and wind turbines. This poses new challenges in the operation of power systems through an increase in the number of generation units and bidirectional power flows [1]. Special challenges arise in power flow optimization, which traditionally relies on known topologies and accurate models [2]. Similarly, accurate power flow calculations have become increasingly important, as operators must respect equipment limits while integrating uncertain renewable sources [3].

Meanwhile, the ongoing digitization of power grids increases the availability of high-resolution data, enabling novel methods. Existing approaches either (i) include an explicit modeling or system identification step or (ii) employ input/output (I/O) data right away to represent systems:

Examples for type (i) are [4] and [5], where power flow is modeled using regression-based methods, mapping node measurements to power flows over power lines. While [4] uses linear mappings, [5] makes use of Koopman operator theory, elevating node measurements into a higher dimensional space where a mapping to power flows is learned.

Type (ii) is based on behavioral systems theory [6]–[8] and eliminates the need for an explicit identification step. Recent studies [9], [10] have employed such approaches for predictive control of power systems. Moreover, in [11], a behavioral approach for controlling multi-energy systems was presented. The methodology incorporated uncertain forecasts of exogenous signals into the data-driven predictive control

formulation. A similar approach was employed in [12] for a data-driven multi-stage DC optimal power flow problem, which requires only measured data of nodal power injection and does not need information about the grid topology. However, DC power flow does not account for transmission losses, rendering the approach less accurate.

To enable applications to nonlinear systems and enhance robustness to noise, the approach presented in [13] introduces regularization into the optimal control problem. Moreover, the authors of [8] extend the application of behavioral systems theory to nonlinear Hammerstein and Wiener systems. However, to the best of our knowledge, such data-driven representations have not yet been employed for nonlinear power flow. We aim to fill this gap by making the following contributions. (i) We derive a data-driven nonlinear active power flow representation for radial networks based on behavioral systems theory that can be used in optimization problems. Our formulation only requires I/O data of power lines, i.e., phase angle differences and transmitted power. (ii) We propose convex relaxations of aforementioned optimization problems, enabling the use of off-the-shelf solvers and convex optimization modeling frameworks. (iii) We demonstrate the applicability of the proposed approaches in a numerical case study on microgrid operation control.

The remainder of this paper is structured as follows. Section II introduces mathematical preliminaries. Sections III and IV introduce physics-based and data-driven power flow optimization problems, respectively. In Section V we use aforementioned formulations in a microgrid example. Section VI presents a numerical case study on this example. Section VII concludes the paper.

## II. MATHEMATICAL PRELIMINARIES

In this section, we cover relevant notations and concepts from data-driven system representations and graph theory.

### A. Notations

We denote real numbers, nonpositive real numbers, and nonnegative real numbers by  $\mathbb{R}$ ,  $\mathbb{R}_{\leq 0}$  and  $\mathbb{R}_{\geq 0}$ , respectively. The sets of positive and nonnegative integers are denoted  $\mathbb{N}$  and  $\mathbb{N}_0$ , respectively. The set of Booleans is  $\mathbb{B} = \{0, 1\}$  and the set of complex numbers is  $\mathbb{C}$ . We express a complex number  $x \in \mathbb{C}$  in polar form as  $x = \hat{x}e^{i\theta}$ , where  $\hat{x} = |x| \in \mathbb{R}$  is the absolute value of  $x$ ,  $\theta \in \mathbb{R}$  is the angle of  $x$ ,  $e$  is Euler’s number, and  $i$  is the imaginary unit. In Cartesian form,  $x$  can be expressed by  $x = \text{Re}(x) + i \text{Im}(x)$ .

A sequence  $\mathbb{W} = \{w(a), w(a+1), \dots, w(b)\}$  is an ordered set of  $b - a + 1$  elements, where  $a, b \in \mathbb{N}_0$ ,  $b \geq a$  and

<sup>1</sup>S. Otzen is with Jouly, Berlin, Germany. [sebastian@jouly.ai](mailto:sebastian@jouly.ai)

<sup>2</sup>H. M. H. Wolf and C. A. Hans are with the Automation and Sensorics in Networked Systems Group, University of Kassel, Germany. [h.wolf@uni-kassel.de](mailto:h.wolf@uni-kassel.de), [hans@uni-kassel.de](mailto:hans@uni-kassel.de)

$w(k) \in \mathbb{R}^{N_w}$  with discrete time instant  $k \in \mathbb{N}_0$ . Set  $\mathbb{W}$  can be compactly written as  $\{w(k)\}_{k=a}^b$ . Throughout this work,  $[w]_{k=a}^b$  is shorthand for  $[w(a)^T \ w(a+1)^T \ \cdots \ w(b)^T]^T$ . Measured values are denoted by superscript  $m$ , e.g.,  $w^m(k)$ . Discrete-time instant  $k$  refers to time  $t = t_0 + kT_s$ , where  $t_0 \in \mathbb{R}_{\geq 0}$  is the initial time and  $T_s \in \mathbb{R}_{\geq 0}$  the sampling time. We use  $w(k+h|k)$  to refer to a prediction for time instant  $k+h$ ,  $h \in \mathbb{N}_0$  that is performed at time  $k$ .

### B. Hankel Matrix

Consider a sequence  $\mathbb{W} = \{w(k)\}_{k=0}^{N-1}$  of length  $N \in \mathbb{N}$ . The associated Hankel matrix of order  $L \in \mathbb{N}$ ,  $L < N$ , is

$$H_L(\mathbb{W}) = \begin{bmatrix} w(0) & w(1) & \cdots & w(N-L) \\ w(1) & w(2) & \cdots & w(N-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ w(L-1) & w(L) & \cdots & w(N-1) \end{bmatrix}. \quad (1)$$

To determine if a sequence allows us to reconstruct the behavior of a system, we recall the following definition.

**Definition 1** (Persistence of Excitation [7]). A sequence  $\mathbb{W} = \{w(k)\}_{k=0}^{N-1}$  with  $w(k) \in \mathbb{R}^{N_w}$  is said to be persistently exciting of order  $L$  if  $H_L(\mathbb{W})$  has full row rank, i.e., if

$$\text{rank}(H_L(\mathbb{W})) = N_w L. \quad (2)$$

### C. Hammerstein Systems

Consider matrices  $A \in \mathbb{R}^{N_x \times N_x}$ ,  $B \in \mathbb{R}^{N_x \times \tilde{N}_u}$ ,  $C \in \mathbb{R}^{N_y \times N_x}$ ,  $D \in \mathbb{R}^{N_y \times \tilde{N}_u}$ , state vector  $x(k) \in \mathbb{R}^{N_x}$ , input  $u(k) \in \mathbb{R}^{N_u}$ , and output  $y(k) \in \mathbb{R}^{N_y}$ . With them, we describe time-invariant Hammerstein systems as [8]

$$x(k+1) = Ax(k) + B\nu(u(k)), \quad x(0) = x_0, \quad (3a)$$

$$y(k) = Cx(k) + D\nu(u(k)), \quad (3b)$$

with the nonlinear mapping  $\nu : \mathbb{R}^{N_u} \rightarrow \mathbb{R}^{\tilde{N}_u}$ . This can be expressed by a linear combination of known basis functions  $\phi_i : \mathbb{R}^{N_u} \rightarrow \mathbb{R}$ ,  $i \in [1, r] \subset \mathbb{N}$  whose span includes  $\nu$ , i.e.,

$$\nu(u(k)) = \sum_{i=1}^r \zeta_i \phi_i(u(k)) = \zeta \phi(u(k)), \quad (4)$$

with  $\zeta = [\zeta_i]_{i=1}^r \in \mathbb{R}^{\tilde{N}_u \times r}$ ,  $\zeta_i \in \mathbb{R}^{\tilde{N}_u}$  not all zero, and

$$\phi(u(k)) = [\phi_1(u(k)) \ \cdots \ \phi_r(u(k))]^T \in \mathbb{R}^r. \quad (5)$$

Combining  $\zeta$  with  $B$  and  $D$  allows us to reformulate (3) as

$$x(k+1) = Ax(k) + \tilde{B}\phi(u(k)), \quad x(0) = x_0, \quad (6a)$$

$$y(k) = Cx(k) + \tilde{D}\phi(u(k)), \quad (6b)$$

with  $\tilde{B} = B\zeta \in \mathbb{R}^{N_x \times r}$  and  $\tilde{D} = D\zeta \in \mathbb{R}^{N_y \times r}$ . In what follows, we will use  $\phi(k)$  as shorthand notation for  $\phi(u(k))$ .

### D. Data-Driven Representation of Hammerstein Systems

**Theorem 1** (Data-driven Hammerstein System [8]). Consider a measured I/O trajectory  $\{u^m(k), y^m(k)\}_{k=n}^{n+N-1}$  of a Hammerstein system (3). Given that the associated auxiliary input  $\{\phi^m(k)\}_{k=n}^{n+N-1}$ , constructed via (5) using  $\{u^m(k)\}_{k=n}^{n+N-1}$ , is persistently exciting of order  $L + N_x$ ,

then,  $\{u(k), y(k)\}_{k=l}^{l+L-1}$  is a trajectory of (3) if and only if there exists an  $\alpha \in \mathbb{R}^{N-L+1}$  such that

$$\begin{bmatrix} H_L(\{\phi^m(k)\}_{k=n}^{n+N-1}) \\ H_L(\{y^m(k)\}_{k=n}^{n+N-1}) \end{bmatrix} \alpha = \begin{bmatrix} [\phi(k)]_{k=l}^{l+L-1} \\ [y(k)]_{k=l}^{l+L-1} \end{bmatrix}, \quad (7a)$$

**Remark 1.** Note that for Theorem 1, the auxiliary input must contain the true basis functions. In practical settings, sufficiently many basis functions or even kernel functions can be used to approximate the true ones [8].

### E. Graph Theory

Following graph-theoretic terminology [14], we define an undirected weighted connected graph  $\mathcal{G} = (\mathbb{V}, \mathbb{E}, y)$  through a set  $\mathbb{V} \subseteq \mathbb{N}_0$  of  $N_b \in \mathbb{N}$  nodes and  $\mathbb{E} \subseteq [\mathbb{V}]^2$  of  $N_e \in \mathbb{N}_0$  edges, where  $[\mathbb{V}]^2$  is the set of all subsets of  $\mathbb{V}$  with two elements. The weighting function is  $y : \mathbb{E} \rightarrow \mathbb{C}^3$  with  $y(\{i, j\}) = (y_{ij}, \bar{y}_{ij}, \bar{y}_{ji})$ . Nodes  $i$  and  $j$  are adjacent if  $\{i, j\} \in \mathbb{E}$ . The set of adjacent nodes of  $i \in \mathbb{V}$  is  $\mathbb{V}_i = \{j \in \mathbb{V} \mid \{i, j\} \in \mathbb{E}\}$ . Let  $\mathbb{E} = \{e_1, e_2, \dots, e_{N_e}\}$  be ordered such that for  $e_l = \{i_l, j_l\}$  and  $e_{l+1} = \{i_{l+1}, j_{l+1}\}$ ,  $l \in [1, N_e - 1] \subset \mathbb{N}$ , it holds that  $i_l \leq i_{l+1}$  and  $j_l \leq j_{l+1}$  if  $i_l = i_{l+1}$ . Moreover, for  $z_{e_l} \in \mathbb{R}^{N_z}$ , we use  $[z_{e_l}]_{\{e_l\} \in \mathbb{E}}$  as shorthand for  $[z_{e_1}^T \ z_{e_2}^T \ \cdots \ z_{e_{N_e}}^T]^T$ .

## III. PHYSICS-BASED POWER FLOW

In this section, we introduce the physics-based active power flow model and define an optimal power flow problem. We model the electrical grid as a weighted undirected connected graph  $\mathcal{G}$ , where each edge  $\{i, j\} \in \mathbb{E}$  represents a power line. We start by posing some assumptions.

**Assumption 1.** The power grid is balanced and symmetrical.

**Assumption 2.** The voltages at all buses are constant, i.e.,  $\hat{v}_i(k) = \hat{v}_i$  for all  $i \in \mathbb{V}$ .

**Assumption 3.** We assume short to medium long transmission lines up to a length of “about 200 km” [15]. Thus, we can model them using the  $\pi$ -equivalent circuit in Figure 1.

#### A. Nonlinear Power Flow Model

Building on Assumption 3, we model connections between two nodes, i.e., edges in  $\mathbb{E}$ , using  $\pi$ -equivalent circuits. The weighting function  $y$  provides series admittance  $y_{ij} = g_{ij} + ib_{ij}$ , with conductance  $g_{ij} \in \mathbb{R}_{\geq 0}$  and susceptance  $b_{ij} \in \mathbb{R}$ , and shunt admittances  $\bar{y}_{ij} = \bar{g}_{ij} + i\bar{b}_{ij}$ ,  $\bar{y}_{ji} = \bar{g}_{ji} + i\bar{b}_{ji}$ .

**Remark 2.** We can also represent more difficult connections between two nodes with our model, e.g., combinations of power lines, transformers and passive loads, since these can be transformed into  $\pi$ -equivalent using Kron-reduction [16].

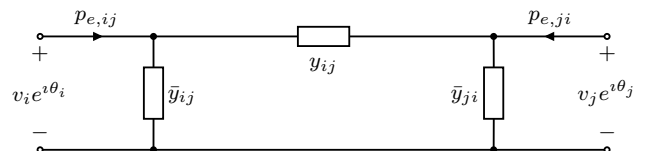


Fig. 1.  $\pi$ -equivalent circuit of a power line.

Consider voltage phasors  $v_i = \hat{v}_i e^{j\theta_i} \in \mathbb{C}$  at all nodes  $i \in \mathbb{V}$ . At time  $k$ , the active power flow from node  $i$  to  $j$  is

$$p_{e,ij}(k) = (\bar{g}_{ij} + g_{ij})\hat{v}_i^2(k) - \hat{v}_i(k)\hat{v}_j(k)(g_{ij}\cos(\theta_{ij}(k)) + b_{ij}\sin(\theta_{ij}(k))), \quad (8)$$

with phase angle difference  $\theta_{ij} = \theta_i - \theta_j$ . Assumption 2 allows us to rewrite (8) as

$$p_{e,ij}(k) = \check{g}_{ij} - (\check{g}_{ij}\cos(\theta_{ij}(k)) + \tilde{b}_{ij}\sin(\theta_{ij}(k))), \quad (9)$$

with  $\check{g}_{ij} = (\bar{g}_{ij} + g_{ij})\hat{v}_i^2$ ,  $\tilde{g}_{ij} = g_{ij}\hat{v}_i\hat{v}_j$ , and  $\tilde{b}_{ij} = b_{ij}\hat{v}_i\hat{v}_j$ .

The net power injection  $p_{g,i} \in \mathbb{R}$ , i.e., the total power provided or consumed at node  $i$ , is equal to the sum of the power flowing into the connected power lines, i.e.,

$$p_{g,i}(k) = \sum_{j \in \mathbb{V}_i} p_{e,ij}(k). \quad (10)$$

### B. Nonlinear Optimal Power Flow

We consider two active power flow values for each power line that we collect in  $p_e(k) = [p_{e,ij}(k) \ p_{e,ji}(k)]_{\{i,j\} \in \mathbb{E}}^T$ . Furthermore, we collect nodal power injections in  $p_g(k) = [p_{g,i}(k)]_{i \in \mathbb{V}}$ . For a power system with  $N_e$  power lines and  $N_b$  nodes we collect line power flows and nodal power injections in  $p(k) = [p_e(k)^T \ p_g(k)^T]^T \in \mathbb{R}^{2N_e+N_b}$  and phase angle differences in  $\theta(k) = [\theta_{ij}(k)]_{\{i,j\} \in \mathbb{E}} \in \mathbb{R}^{N_e}$ . Let us now employ (9) and (10) in an optimization problem.

**Problem 1** (Nonlinear optimal power flow).

$$\min_{\theta(k), p(k)} f(p(k))$$

subject to (10)  $\forall i \in \mathbb{V}$  and

$$p_{e,ij}(k) = \check{g}_{ij} - (\check{g}_{ij}\cos(\theta_{ij}(k)) + \tilde{b}_{ij}\sin(\theta_{ij}(k))),$$

$$p_{e,ji}(k) = \check{g}_{ji} - (\check{g}_{ji}\cos(-\theta_{ij}(k)) + \tilde{b}_{ij}\sin(-\theta_{ij}(k))),$$

$\forall \{i, j\} \in \mathbb{E}$  as well as application-dependent constraints

$$0 \leq f_{\text{ineq}}(\theta(k), p(k)), \quad (12a)$$

$$0 = f_{\text{eq}}(\theta(k), p(k)). \quad (12b)$$

Here, the objective  $f$  describes the cost of operation, e.g., in the form of generation cost or the transmission losses

$$p_{\text{loss}}(k) = \sum_{i \in \mathbb{V}} p_{g,i}(k). \quad (13)$$

## IV. DATA-DRIVEN POWER FLOW

In what follows, we transform the physics-based model into a Hammerstein system using appropriate basis functions. We then apply Theorem 1 and employ the resulting data-driven active power flow into an optimization problem. We again start with some assumptions.

**Assumption 4.** We consider radial networks, which are widely present in microgrid settings [17]. As discussed in Remark 6, this leads to unique power flow solutions and enables exact convex relaxations.

**Assumption 5.** Voltage phase angles  $\theta_{ij}(k)$  are available for all lines, e.g., through time-synchronized phasor measurement units (PMUs) or micro-PMUs. Moreover, we assume that active power measurements of all lines are available.

### A. Data-Driven Representation of Nonlinear Power Flow

Equation (9) for  $\{i, j\} \in \mathbb{E}$  can be understood as a stateless I/O system with input  $u(k) = \theta_{ij}(k)$  and output  $y(k) = p_{e,ij}(k)$ . We can transform it into a Hammerstein system (6) using the basis functions 1,  $\cos(\theta_{ij}(k))$ , and  $\sin(\theta_{ij}(k))$ , i.e.,

$$\phi(k) = [1 \ \cos(\theta_{ij}(k)) \ \sin(\theta_{ij}(k))]^T. \quad (14)$$

This allows us to express (9) as

$$p_{e,ij}(k) = [\check{g}_{ij} \ -\tilde{g}_{ij} \ -\tilde{b}_{ij}]\phi(k). \quad (15)$$

By virtue of Theorem 1, this can be represented in a data-driven manner as

$$\begin{bmatrix} H_1(\{\phi^m(k)\}_{k=n}^{n+N-1}) \\ H_1(\{p_{e,ij}^m(k)\}_{k=n}^{n+N-1}) \end{bmatrix} \alpha(k) = \begin{bmatrix} \phi(k) \\ p_{e,ij}(k) \end{bmatrix}. \quad (16)$$

**Remark 3.** Note that we "consider algebraic relations as differential equations of order zero" [6, p. 12], i.e.,  $N_x = 0$ ,  $A = 0$ ,  $\tilde{B} = 0$ ,  $C = 0$ , and  $\tilde{D} \neq 0$ . Consequently,  $L = 1$ .

### B. Data-Driven Optimal Power Flow

To employ (16) in optimization problems, we extend (14) by trigonometric basis functions for all power lines, i.e.,

$$\phi(k) = [1 \ \cos(\theta_{ij}(k)) \ \sin(\theta_{ij}(k))]_{\{i,j\} \in \mathbb{E}}^T. \quad (17)$$

This allows us to derive a data-driven representation of the nonlinear active power flow for the entire grid using  $\phi^m(k)$  which is constructed along the lines of (17) and  $p_e^m(k)$ .

**Remark 4.** Due to the symmetry properties,  $\sin(\theta_{ji}(k))$  and  $\cos(\theta_{ji}(k))$  do not need to be explicitly included.

In optimization problems, we treat  $\phi(k)$  as a vector of decision variables. To maintain the inherent relation between the trigonometric components, we enforce  $\cos^2(\theta_{ij}(k)) + \sin^2(\theta_{ij}(k)) = 1$  for all  $\{i, j\} \in \mathbb{E}$ , i.e.,  $\phi_l^2(k) + \phi_{l+1}^2(k) = 1$  for  $l \in \{2, 4, \dots, 2N_e\}$ . With  $\alpha(k) \in \mathbb{R}^N$ , we formulate the following optimization problem.

**Problem 2** (Data-driven optimal power flow).

$$\min_{\alpha(k), \phi(k), p(k)} f(p(k))$$

subject to (10)  $\forall i \in \mathbb{V}$ , (12) and

$$\begin{bmatrix} H_1(\{\phi^m(k)\}_{k=n}^{n+N-1}) \\ H_1(\{p_e^m(k)\}_{k=n}^{n+N-1}) \end{bmatrix} \alpha(k) = \begin{bmatrix} \phi(k) \\ p_e(k) \end{bmatrix}, \quad (18a)$$

$$\phi_l^2(k) + \phi_{l+1}^2(k) = 1, \quad \forall l \in \{2, 4, \dots, 2N_e\}. \quad (18b)$$

**Remark 5.** For a radial network with  $N_e = N_b - 1$  lines and  $N_b$  nodes, Problem 2 introduces  $\phi(k) \in \mathbb{R}^{2N_e+1}$  and  $\alpha(k) \in \mathbb{R}^N$  as decision variables. Since Definition 1 requires  $N \geq 2N_e + 1 = 2N_b - 1$ , the total number of decision variables and constraints grows linearly in  $N_b$ .

Due to (18b), Problem 2 is nonconvex. Such problems are computationally demanding and finding a global optimum can be hard. This issue is addressed in the next section.

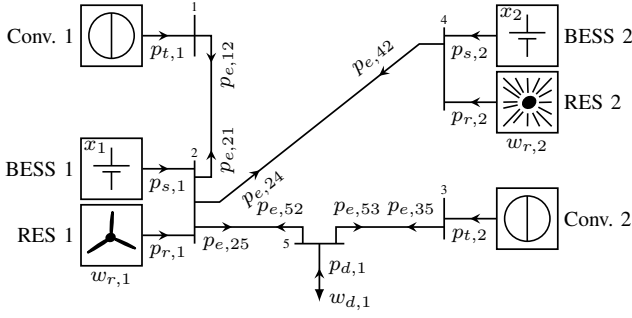


Fig. 2. Example microgrid.

### C. Convex Data-Driven Optimal Power Flow

Convex formulations guarantee convergence to a globally optimal solution and improved scalability. Inspired by [18], we therefore propose a convex relaxation that replaces (18b) with the unit disk (19). To promote solutions on the unit disk boundary, we add a regularization term that maximizes cosine basis functions. The problem reads as follows.

**Problem 3** (Convex data-driven optimal power flow).

$$\min_{\alpha(k), \phi(k), p(k)} f(p(k)) - \beta \sum_{i=1}^{N_e} \phi_{2i}(k)$$

subject to (10)  $\forall i \in \mathbb{V}$ , (12), (18a) and

$$\phi_l^2(k) + \phi_{l+1}^2(k) \leq 1, \quad \forall l \in \{2, 4, \dots, 2N_e\}. \quad (19)$$

**Remark 6.** For radial networks, replacing (18b) by (19) is exact if (C1) the cost function  $f(p)$  is strictly increasing in all nodal power injections and (C2) the phase angle differences are sufficiently small [19, Theorem 8]. In light of Assumption 3, (C2) is typically satisfied. If (C1) holds,  $\beta$  can be set to zero. Otherwise, a small  $\beta > 0$  promotes solutions on the unit circle. The parameter  $\beta$  should be chosen such that the resulting cost distortion  $\beta N_e \theta_{\max}^2/2$ , where  $\theta_{\max}$  is the largest phase angle difference, remains negligible.

## V. MICROGRID APPLICATION EXAMPLE

In this section, we present a microgrid (MG) application example that uses the proposed methods in a model predictive control (MPC) scheme. While our approach is applicable to large radial networks (see Remark 5), we use an MG due to its illustrative character. The employed setup in Figure 2 is based on [20] and composed of 5 buses  $\mathbb{V} = \{1, 2, 3, 4, 5\}$  as well as 4 power lines, i.e.,  $\mathbb{E} = \{\{1, 2\}, \{2, 4\}, \{2, 5\}, \{3, 5\}\}$ . Furthermore, it includes two conventional units, two battery energy storage systems (BESSs), two renewable energy sources (RES) and a load.

### A. Constraints

We collect the output power of the conventional generators in  $p_t(k) = [p_{t,1}(k) \ p_{t,2}(k)]^T \in \mathbb{R}_{\geq 0}^2$  and their switch condition in  $\delta_t(k) = [\delta_{t,1}(k) \ \delta_{t,2}(k)]^T \in \mathbb{B}^2$  where  $\delta_{t,i}(k) = 1$  means that generator  $i \in \{1, 2\}$  is enabled and  $\delta_{t,i}(k) = 0$  that it is disabled. When enabled, the generators operate

within their power limits,  $p_t^{\min}, p_t^{\max} \in \mathbb{R}_{\geq 0}^2$ , when disabled their respective power is zero. This is captured by

$$\text{diag}(p_t^{\min}) \delta_t(k) \leq p_t(k) \leq \text{diag}(p_t^{\max}) \delta_t(k). \quad (20)$$

The BESSs can either draw power from the grid or supply power to it. The BESSs' power  $p_s = [p_{s,1} \ p_{s,2}]^T \in \mathbb{R}^2$  must remain within the limits  $p_s^{\min} \in \mathbb{R}_{\leq 0}^2, p_s^{\max} \in \mathbb{R}_{\geq 0}^2$ , i.e.,

$$p_s^{\min} \leq p_s(k) \leq p_s^{\max}. \quad (21)$$

The stored energy of the BESSs  $x \in \mathbb{R}_{\geq 0}^2$  evolves over time based on  $p_s$  and the initially stored energy  $x_0 \in \mathbb{R}_{\geq 0}^2$ , i.e.,

$$x(k+1) = A_s x(k) + B_s p_s(k), \quad x(0) = x_0, \quad (22)$$

where  $A_s \in \mathbb{R}^{2 \times 2}$  and  $B_s \in \mathbb{R}^{2 \times 2}$  are diagonal matrices. The stored energy is bounded by  $x^{\min}, x^{\max} \in \mathbb{R}_{\geq 0}^2$ , i.e.,

$$x^{\min} \leq x(k) \leq x^{\max}. \quad (23)$$

RES provide power depending on wind or sunlight. Their infeed  $p_r(k) = [p_{r,1}(k) \ p_{r,2}(k)]^T \in \mathbb{R}_{\geq 0}^2$  can not exceed the weather-dependent available power,  $w_r(k) \in \mathbb{R}_{\geq 0}^2$ , i.e.,

$$0 \leq p_r(k) \leq w_r(k). \quad (24)$$

The active power flowing through power lines is limited by  $p_e^{\min} \in \mathbb{R}_{\leq 0}^8$  and  $p_e^{\max} \in \mathbb{R}_{\geq 0}^8$ , i.e.,

$$p_e^{\min} \leq p_e(k) \leq p_e^{\max}. \quad (25)$$

The load demand is  $p_d(k) \in \mathbb{R}_{\leq 0}$ . The consumed or provided power at the nodes depends on the connected units/loads, i.e.,

$$p_g(k) = U[p_t(k)^T \ p_s(k)^T \ p_r(k)^T \ p_d(k)^T]^T \quad (26)$$

where  $U \in \mathbb{B}^{5 \times 7}$  is a matrix with entries

$$U_{ij} = \begin{cases} 1, & \text{if unit/load } j \text{ is connected to node } i, \\ 0, & \text{otherwise.} \end{cases}$$

### B. Cost

The cost  $\ell \in \mathbb{R}$  to operate the MG at time instant  $k$  is

$$\ell(k) = \ell_t^{sw}(k) + \ell_p(k) + \ell_x(k) + \ell_{loss}(k). \quad (27)$$

With  $c_0, c_1 \in \mathbb{R}_{\geq 0}^2$ , the costs of enabling/disabling conventional units plus power-independent running costs are

$$\ell_t^{sw}(k) = c_0^T |\delta_t(k) - \delta_t(k-1)| + c_1^T \delta_t(k), \quad (28)$$

The cost  $\ell_p$  models fuel costs, incentivizes RES infeed and accounts for storage conversion losses, i.e.,

$$\ell_p(k) = c_2^T p_t(k) + c_3^T p_r(k) + c_4^T |p_s(k)|, \quad (29)$$

with  $c_2 \in \mathbb{R}_{\geq 0}^2, c_3 \in \mathbb{R}_{\geq 0}^2$  and  $c_4 \in \mathbb{R}_{\geq 0}^2$ .

To penalize deviations of the stored energy  $x(k)$  outside the permissible range  $[\underline{x}, \bar{x}]$ , we use  $c_5 \in \mathbb{R}_{\geq 0}^2$  and formulate

$$\ell_x(k) = c_5^T (\max(0, \underline{x} - x(k)) + \max(0, x(k) - \bar{x})), \quad (30)$$

Finally, the cost  $\ell_{loss}$  is modeled with  $c_6 \in \mathbb{R}_{\geq 0}$  as

$$\ell_{loss}(k) = c_6 p_{loss}(k). \quad (31)$$

### C. Model Predictive Control Problems

The decision variables of the MPC problems are

$$P = \left[ [p_u(k+h|k)^T \ p_g(k+h|k)^T \ p_e(k+h|k)^T]^T \right]_{h=0}^{H-1}$$

where  $p_u(k) = [p_t(k)^T \ p_s(k)^T \ p_r(k)^T]^T$ , and  $H-1$  is the prediction horizon. We further define  $\Delta_t = [\delta_t(k+h|k)]_{h=0}^{H-1}$ ,  $X = [x(k+h|k)]_{h=0}^{H-1}$ , and  $\Theta = [\theta(k+h|k)]_{h=0}^{H-1}$ . Finally, discount factor  $\gamma \in (0, 1)$  is used to prioritize near-future decisions and formulate the following MPC problem.

**Problem 4** (Nonlinear MPC for MG operation).

$$\min_{\Theta, P, \Delta_t, X} \sum_{h=0}^{H-1} \ell(k+h|k) \gamma^h$$

subject to (20)–(26) as well as

$$p_{g,i}(k+h|k) = \sum_{j \in \mathbb{V}_i} p_{e,ij}(k+h|k), \quad \forall i \in \mathbb{V}, \quad (32a)$$

$$p_{e,ij}(k+h|k) = \check{g}_{ij} - (\check{g}_{ij} \cos(\theta_{ij}(k+h|k)) + \tilde{b}_{ij} \sin(\theta_{ij}(k+h|k))), \quad \forall \{i, j\} \in \mathbb{E}, \quad (32b)$$

$$p_{e,ji}(k+h|k) = \check{g}_{ji} - (\check{g}_{ji} \cos(-\theta_{ij}(k+h|k)) + \tilde{b}_{ij} \sin(-\theta_{ij}(k+h|k))), \quad \forall \{i, j\} \in \mathbb{E}, \quad (32c)$$

for  $h = 0, \dots, H-1$  with initially available  $x(k|k) = x_k$  and  $\delta_t(k-1|k) = \delta_{t,k-1}$ .

Problem 4 is a nonconvex, mixed-integer optimization problem with an explicit model of the power lines assuming known parameters. In what follows, we employ the data-driven representations from Section IV. Therefore, we use  $\Phi = [\phi(k+h|k)]_{h=0}^{H-1}$  with  $\phi(k)$  from (17) and  $\alpha = [\alpha(k+h|k)]_{h=0}^{H-1}$ ,  $\alpha(k) \in \mathbb{R}^9$  in the following problem.

**Problem 5** (Nonconvex data-driven MPC).

$$\min_{\alpha, \Phi, P, \Delta_t, X} \sum_{h=0}^{H-1} \ell(k+h|k) \gamma^h$$

subject to (20)–(26) and (32a) as well as

$$\begin{bmatrix} H_1(\{\phi^m(k)\}_{k=n}^{n+N-1}) \\ H_1(\{p_e^m(k)\}_{k=n}^{n+N-1}) \end{bmatrix} \alpha(k+h|k) = \begin{bmatrix} \phi(k+h|k) \\ p_e(k+h|k) \end{bmatrix}, \quad (33a)$$

$$\phi_i^2(k+h|k) + \phi_{i+1}^2(k+h|k) = 1, \quad \forall i \in [2, 4, 6, 8], \quad (33b)$$

for  $h = 0, \dots, H-1$  with  $x(k|k) = x_k$ ,  $\delta_t(k-1|k) = \delta_{t,k-1}$ .

Problem 5 is nonconvex due to (33b). To enable the use of off-the-shelf solvers and convex optimization formulation frameworks, we employ the relaxation from Section IV-C.

**Problem 6** (Convex data-driven MPC).

$$\min_{\alpha, \Phi, P, \Delta_t, X} \sum_{h=0}^{H-1} (\ell(k+h|k) \gamma^h - \beta \sum_{i \in \{2, 4, 6, 8\}} \phi_i(k+h|k))$$

subject to (20)–(26), (32a) and (33a) as well as

$$\phi_i^2(k+h|k) + \phi_{i+1}^2(k+h|k) \leq 1, \quad \forall i \in [2, 4, 6, 8], \quad (34)$$

for  $h = 0, \dots, H-1$  with  $x(k|k) = x_k$ ,  $\delta_t(k-1|k) = \delta_{t,k-1}$ .

## VI. CASE STUDY

We will now evaluate the closed-loop performance of different MPC approaches in a numerical case study. First, we describe the simulation setup and the parameters. Then, we assess the performance over a seven day simulation.

### A. Simulation Setup

We implemented the case study in Python. The MPC problems were formulated using CVXPY [21] for convex and Gurobi for nonconvex problems. All problems were solved using Gurobi 12, which employs different techniques to find the global optima [22]. Hankel matrices were constructed offline from a prior simulation of the physics-based model excited with persistently exciting inputs and kept fixed throughout the case study. In this context, we used  $N = 9$  samples, the minimum required by Remark 5. All computations were executed on an Intel Core i7-8565U CPU with 16 GB RAM. Simulations span seven days at  $T_s = 0.5$  h ( $K = 336$  steps). An ideal forecaster with perfect knowledge about the future was used to focus on effects associated with the different active power flow representations. All parameters are given in Table I. Here, power line data was taken from [23] while all remaining parameters were taken from [20].

**Remark 7.** Since  $c_{3,2} + c_6 = 0$  for the second RES, the cost function is nondecreasing but not strictly increasing, so Condition (C1) of Remark 6 does not hold. We therefore set  $\beta = 1$ . Since  $\theta_{\max} \leq 0.05$  rad, the cost distortion bound evaluates to  $\beta N_e \theta_{\max}^2 / 2 \leq 0.005$ , which is small compared to the average optimal cost.

### B. Results

Figure 3 shows the results of simulations with Problem 5. Simulations for Problems 4 and 6 were also performed. Within error bounds of  $10^{-4}$ , all formulations yielded identical trajectories of all variables, including the power flow, and identical objective values  $\bar{\ell}_o = \frac{1}{K} \sum_{k=1}^K (\ell_t^{sw}(k) + \ell_p(k)) = -0.2602$  and  $\bar{\ell}_{loss} = \frac{1}{K} \sum_{k=1}^K \ell_{loss}(k) = 0.0047$ , while all constraints were satisfied. This indicates that the data-driven formulations can successfully replace the physics-based models and that the  $\beta = 1$  regularization introduces negligible cost distortion (see Remark 6).

TABLE I  
PARAMETERS OF THE MG MODEL [20].

| Param.          | Value                                           | Param.       | Value                                           | Param.          | Value                                                              |
|-----------------|-------------------------------------------------|--------------|-------------------------------------------------|-----------------|--------------------------------------------------------------------|
| $c_0$           | $\begin{bmatrix} 0.2 \\ 0.1 \end{bmatrix}$      | $c_1$        | $\begin{bmatrix} 0.13 \\ 0.07 \end{bmatrix}$    | $c_2$           | $\begin{bmatrix} 1.56 \\ 1.43 \end{bmatrix}$                       |
| $c_3$           | $\begin{bmatrix} -0.8 \\ -1.0 \end{bmatrix}$    | $c_4$        | $\begin{bmatrix} 0.1 \\ 0.05 \end{bmatrix}$     | $c_5$           | $\begin{bmatrix} 10^3 \\ 10^3 \end{bmatrix}$                       |
| $c_6$           | 1                                               | $\gamma$     | 0.9                                             | $g_{ij}$        | 2 pu                                                               |
| $b_{ij}$        | -20 pu                                          | $p_t^{\min}$ | $\begin{bmatrix} 0.3 \\ 0.1 \end{bmatrix}$ pu   | $p_t^{\max}$    | $\begin{bmatrix} 0.9 \\ 0.6 \end{bmatrix}$ pu                      |
| $p_e^{\min}$    | $[-1 N_e]$ pu                                   | $p_s^{\min}$ | $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$ pu     | $p_s^{\max}$    | $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ pu                          |
| $p_e^{\max}$    | $[1 N_e]$ pu                                    | $A_s$        | $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  | $B_s$           | $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ |
| $x^{\min}$      | $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ pu h     | $x^{\max}$   | $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$ pu h     | $x_0$           | $\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$ pu h                    |
| $\underline{x}$ | $\begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$ pu h | $\bar{x}$    | $\begin{bmatrix} 6.5 \\ 3.5 \end{bmatrix}$ pu h | $\delta_{0,-1}$ | $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$                             |
| $H$             | 6                                               | $\beta$      | 1                                               |                 |                                                                    |

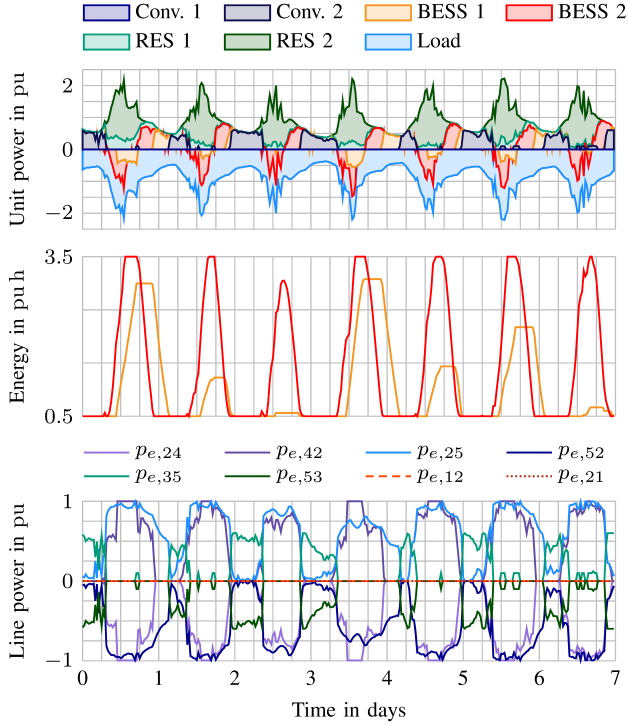


Fig. 3. Simulation results of controlling the reference MG using the convex data-driven MPC over seven days. The unit power is presented as a stacked area plot.

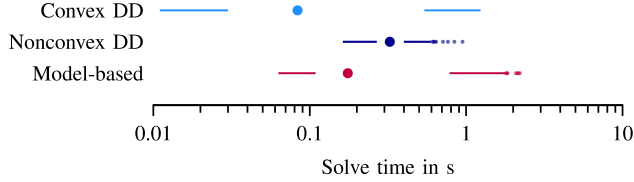


Fig. 4. Solve times across all 336 time steps for the model-based Problem 4, the nonconvex data-driven Problem 5, and the convex data-driven Problem 6. The large dots are the median, the surrounding white area the interquartile range, and the whiskers extend to values within 1.5 times the interquartile range. Values beyond the whiskers are outliers.

Figure 4 shows the distribution of the solve times for all 336 time steps. The convex data-driven MPC achieves the shortest median solve time of 0.08s, followed by the reference MPC at 0.18s and the nonconvex data-driven MPC at 0.33s, which has the highest median solve time. Given a sampling time of 30 minutes, the maximum solve times are well within the real-time requirements of the system at hand, demonstrating the applicability of the proposed approaches.

## VII. CONCLUSIONS

We proposed data-driven formulations of nonlinear active power flows based on behavioral systems theory. Our approaches enable active power flow optimization without relying on physics-based power line models and without requiring an explicit system identification step. Within error bounds of  $10^{-4}$ , all approaches yielded identical trajectories and objective values while satisfying all constraints. Moreover, all approaches had acceptable solve times below 3s.

Future research will evaluate scalability, extensions to meshed networks and topology-agnostic formulations, and robustness to measurement noise.

## REFERENCES

- [1] C. Bordons, *Model predictive control of microgrids*, ser. Advances in industrial control, F. Garcia-Torres and M. A. Ridao, Eds. Springer, 2020.
- [2] S. H. Low, “Convex relaxation of optimal power flow—part i: Formulations and equivalence,” *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 15–27, 2014.
- [3] N. Yu, S. Zhang, J. Qin, P. Hidalgo-Gonzalez, R. Dobbe, Y. Liu, A. Dubey, Y. Wang, J. Dirkman, H. Zhong, N. Lu, E. Ma, Z. Ding, D. Cao, J. Zhao, and Y. Gao, “Data-driven control, optimization, and decision-making in active power distribution networks,” *Applied Energy*, vol. 397, p. 126253, 2025.
- [4] Y. Liu, N. Zhang, Y. Wang, J. Yang, and C. Kang, “Data-driven power flow linearization: A regression approach,” *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2569–2580, 2019.
- [5] L. Guo, Y. Zhang, X. Li, Z. Wang, Y. Liu, L. Bai, and C. Wang, “Data-driven power flow calculation method: A lifting dimension linear regression approach,” *IEEE Trans. Power Syst.*, vol. 37, no. 3, pp. 1798–1808, 2022.
- [6] J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory*. Springer, 1998.
- [7] J. Willems, I. Markovsky, P. Rapisarda, and B. De Moor, “A note on persistency of excitation,” in *43rd CDC*, 2004.
- [8] J. Berberich and F. Allgöwer, “A trajectory-based framework for data-driven system analysis and control,” in *ECC*, 2020, pp. 1365–1370.
- [9] I. Markovsky, L. Huang, and F. Dörfler, “Data-driven control based on the behavioral approach: From theory to applications in power systems,” *IEEE Control Syst.*, vol. 43, no. 5, pp. 28–68, 2023.
- [10] J. B. Lipka and C. A. Hans, “Data-driven model predictive control of battery storage units,” in *ACC*, 2024, pp. 2230–2235.
- [11] D. Bilgic, A. Koch, G. Pan, and T. Faulwasser, “Toward data-driven predictive control of multi-energy distribution systems,” *Electric Power Systems Research*, vol. 212, p. 108311, 2022.
- [12] O. Molodchyk, P. Schmitz, A. Engelmann, K. Worthmann, and T. Faulwasser, “Towards data-driven multi-stage opf,” in *2025 IEEE Kiel PowerTech*. IEEE, 2025, pp. 1–6.
- [13] J. Coulson, J. Lygeros, and F. Dörfler, “Data-enabled predictive control: In the shallows of the deep,” in *ECC*, 2019.
- [14] R. Diestel, *Graph Theory*, fifth edition ed., ser. Graduate texts in mathematics. Berlin: Springer, 2017, no. 173.
- [15] P. Kundur, *Power system stability and control*. McGraw-Hill, 2007.
- [16] F. Dörfler and F. Bullo, “Kron Reduction of Graphs With Applications to Electrical Networks,” *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 60, no. 1, pp. 150–163, 2013.
- [17] R. Lasseter, “Microgrids,” in *IEEE PES WM*, 2002, pp. 305–308.
- [18] R. Jabr, “Radial distribution load flow using conic programming,” *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1458–1459, 2006.
- [19] S. H. Low, “Convex relaxation of optimal power flow—part II: Exactness,” *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 2, pp. 177–189, Jun. 2014.
- [20] C. A. Hans, *Operation control of islanded microgrids*. Shaker Verlag, 2021.
- [21] S. Diamond and S. Boyd, “CVXPY: A Python-embedded modeling language for convex optimization,” *J. Mach. Learn. Res.*, vol. 17, no. 83, pp. 1–5, 2016.
- [22] Gurobi Optimization, Inc. (2026) Gurobi optimizer: Nonlinear features. [Online]. Available: <https://docs.gurobi.com/projects/optimizer/en/current/features/nonlinear.html>
- [23] H. M. H. Wolf and C. A. Hans, “Identification of power systems with droop-controlled units using neural ordinary differential equations,” in *ECC*, 2025, pp. 800–806.