Electric Power Systems Research 235 (2024) 110816

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Distributed sequential optimal power flow under uncertainty in power distribution systems: A data-driven approach ☆

Georgios Tsaousoglou [a,*,1], Petros Ellinas [b], Juan S. Giraldo [c,1], Emmanouel Varvarigos [b]

[a] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Denmark
[b] Institute of Communications and Computer Systems, National Technical University of Athens, Greece
[c] The Netherlands Organisation for Applied Scientific Research (TNO), The Netherlands

ARTICLE INFO

Keywords:
Optimal power flow
Uncertainty
Optimal control
Data-driven optimization
Sequential decisions

ABSTRACT

Modern distribution systems with high penetration of distributed energy resources face multiple sources of uncertainty. This transforms the traditional Optimal Power Flow problem into a problem of sequential decision-making under uncertainty. In this framework, the solution concept takes the form of a *policy*, i.e., a method of making dispatch decisions when presented with a real-time system state. Reasoning over the future uncertainty realization and the optimal online dispatch decisions is especially challenging when the number of resources increases and only a small dataset is available for the system's random variables. In this paper, we present a data-driven distributed policy for making dispatch decisions online and under uncertainty. The policy is assisted by a Graph Neural Network but is constructed in such a way that the resulting dispatch is guaranteed to satisfy the system's constraints. The proposed policy is experimentally shown to achieve a performance close to the optimal-in-hindsight solution, significantly outperforming state-of-the-art policies based on stochastic programming and plain machine-learning approaches.

1. Introduction

1.1. Motivation

The growing penetration of distributed energy resources (DERs) constitutes a cornerstone development of modern power systems towards supporting higher levels of renewable energy and system flexibility. This development creates, however, significant challenges for Distribution System Operators (DSOs) — the entities responsible for maintaining the system's operation within safe technical limits in an economically efficient way. This predominantly refers to solving the renowned Optimal Power Flow (OPF) problem.

An important challenge refers to the increasingly high levels of uncertainty which motivates solving the OPF problem in a stochastic and adaptive fashion. This brings the standard OPF problem into the realm of sequential decision-making under uncertainty. At the same time, the diversity of DERs impedes solutions based on comprehensive modeling approaches, while their multitude and their distributed nature makes it difficult to manage them centrally. These challenges motivate data-driven and distributed decision-making approaches, as elaborated in [1,2], respectively.

1.2. Related work

The previous subsection motivated the consideration of a data-driven, distributed, sequential, and uncertainty-aware solution to the OPF problem. In this subsection, we discuss the related literature with respect to these requirements.

Considering the constraint-aware economic dispatch of DERs for a look-ahead horizon, the simplest approach is to model uncertain parameters using point-forecast estimations of their future values and solve a deterministic OPF problem for the horizon. Such an optimization problem can be readily extended to a distributed optimization counterpart [2], using decomposition (e.g. [3]) approaches. The distributed optimization approach for the second-order cone (SOCP) formulation in particular, is analyzed in [4]. Such deterministic-optimization-based approaches can readily form the component of an adaptive (rolling-horizon) algorithm, where the deterministic look-ahead optimization is

☆ This work was partly funded by the Horizon 2020 research and innovation programme under the ARV project (grant agreement no. 101036723).
* Corresponding author.
E-mail address: geots@dtu.dk (G. Tsaousoglou).
1 Member, IEEE.

https://doi.org/10.1016/j.epsr.2024.110816
Received 9 August 2023; Received in revised form 22 February 2024; Accepted 17 June 2024
Available online 28 June 2024

Nomenclature

A. Sets

$\mathcal{N}$	Set of energy resources.
$\mathcal{T}$	Set of operating timeslots.
$\mathcal{B}$	Set of distribution network nodes/buses.
$\mathcal{C}_b$	Set of descendant (children) nodes of node b.
$\mathcal{T}$	Index and set of operating timeslots.
$\mathcal{E}_n$	Set of EV's (random) characteristics.
$\mathcal{U}$	Set of decision/control variables.
$\mathcal{W}_t$	Set of random variables for timeslot t.
$\mathcal{S}_t$	Set of state variables for timeslot t.
$\mathcal{F}_t$	Set of belief-state parameters for timeslot t.
$\mathcal{K}$	Set of scenarios for the stochastic program.
$\mathcal{I}$	Set of neural network inputs.
$\mathcal{O}$	Set of neural network outputs.

B. Parameters

$\overline{\mathrm{P}}_n, \underline{\mathrm{P}}_n$	Resource's upper/lower bounds on energy generation or consumption.
$\mathrm{RU}_n, \mathrm{RD}_n$	Generator's ramp up/down.
g_n	Generator's per-unit fuel cost.
$\overline{\mathrm{E}}_n, \underline{\mathrm{E}}_n$	Battery's upper/lower bound on energy containment.
$\overline{\mathrm{V}}, \underline{\mathrm{V}}$	Upper/lower bound on voltages.
$\bar{\mathrm{I}}_{\zeta_b b}$	Upper bound on line's current.
$\lambda_{b,t}$	Lagrange multiplier of the active power balance constraint.
$\mu_{b,t}$	Lagrange multiplier of the reactive power balance constraint.
ρ	Tuning parameter of the ADMM algorithm.

C. Decision Variables

$x_{n,t}$	Resource's dispatch control variable.
$Q_{n,t}$	Resource's reactive power injection at t.
y_n	Energy not delivered to an EV.
$P_{ij,t}$	Active power flow between nodes i and j at t.
$Q_{ij,t}$	Reactive power flow between nodes i and j at t.
$I^{\mathrm{sqr}}_{bc,t}$	Squared magnitude of current flowing between nodes b and c at t.
$V^{\mathrm{sqr}}_{b,t}$	Squared magnitude of node voltage at t.

C. Random Variables

$\widetilde{\mathrm{G}}_{n,t}$	RES generation at t.
$\widetilde{\mathrm{t}}^{\mathrm{arr}}_n$	EV's arrival time.
$\widetilde{\mathrm{t}}^{\mathrm{dep}}_n$	EV's departure time.
$\widetilde{\mathrm{E}}^{\mathrm{arr}}_n$	EV's initial state of charge.
$\widetilde{\mathrm{C}}_n$	EV's battery capacity.
$\widetilde{\mathrm{E}}^{\mathrm{des}}_n$	EV's desired state of charge at departure.
$\widetilde{\mathrm{p}}_t$	Wholesale electricity price at t.
$\widetilde{\mathrm{D}}_{n,t}$	Consumer demand at t.

re-solved at each decision stage using updated estimations for uncertain parameters, as in [5].

Naturally, the point-forecast optimization reduces all the statistical knowledge about an uncertain parameter to a single value (namely, the parameter's expected value). Such a simplistic approach can have detrimental effects to the solution's efficiency and recent literature has proposed more sophisticated methods for solving the OPF problem under uncertainty. Indicatively, [6] proposed a distributed scenario-based stochastic programming approach for the SOCP model. Furthermore, the authors in [7] used the Markov Decision Process (MDP) framework to model the problem of minimizing the distribution system's expected operational cost under network constraints and presented an approximate dynamic programming approach for approximating the optimal solution. The authors in [8] presented distributed solutions where each resource solves its local MDP and the DSO receives the responses and updates a set of Lagrange multipliers, similarly to the above-mentioned distributed optimization techniques.

Notably, the methods reviewed so far are model-based, in the sense that the system's uncertain parameters are assumed to follow known statistical models and/or their temporal dynamics are assumed to follow known transition functions, while some of the methods can also be computationally intensive which puts their suitability for the online adaptive OPF problem into question. In contrast, data-driven approaches refrain from making distributional assumptions about the system's random variables while the relevant Machine Learning (ML) techniques are able to make fast dispatch decisions online, once presented with the information about the system's current state; this is also referred to as the "learn-to-optimize" concept (see [9] for an extensive analysis, and [10] for its application to a AC-OPF). Recently introduced methodological enhancements, tailored to the OPF problem, include the co-called physics-informed neural networks [11] and sensitivity-informed neural networks [12]. The main issue with ML-based methods, however, is that they generally lack constraint-satisfaction guarantees.

1.3. Research gap & contributions

The literature review reveals a number of requirements for an operational policy that makes dispatch decisions in an active distribution network. Namely, the decisions are to be made:

(1) stochastically, i.e., in an uncertainty-aware manner;
(2) adaptively, i.e. in a sequential manner, each time accounting for the updated information;
(3) in a data-driven manner to avoid making statistical assumptions about the uncertainties;
(4) distributedly, for scalability and privacy preservation;
(5) reliably, i.e., in a way that guarantees the satisfaction of network constraints.

In this paper, we first formulate the relevant problem of constructing an optimal policy, using the unified framework for sequential decisions proposed in [13] and the SOCP-relaxation of the OPF. After presenting two benchmark policies, one based on stochastic programming and one based on the learning-to-optimize approach, we proceed to construct the proposed policy by training a Graph Neural Network to estimate the optimal dual variables of the system's power balance constraints for the future stages of the look-ahead horizon. The Graph Neural Network is able to leverage the spatial dependencies of the distribution system to optimize its estimation for the system's optimal duals. Using these estimations as uncertainty-capturing signals, we employ a distributed optimization algorithm that converges to constraint-satisfying here-and-now dispatch decisions. Thereby, the paper's contributions can be summarized as follows:

- A data-driven, distributed policy is presented for the stochastic sequential OPF problem, which makes sure that the system's constraints are respected.
- The proposed policy is shown to compare favorably against stochastic programming and plain ML methods.

2. System model

2.1. Distributed energy resources

We consider a set $\mathcal{N}$ of flexible electricity consuming/producing resources. Each resource belongs to a particular set $\mathcal{N}_\theta$ of type $\theta \in \Theta$, where

$\Theta = \{\text{Generators}, \text{RES}, \text{Consumers}, \text{Storage}, \text{EVs}\}$

is the set of types. It is $\mathcal{N} = \bigcup_{\theta\in\Theta} \mathcal{N}_\theta$. Continuous time is divided into timeslots of equal duration for a time horizon $\mathcal{T}$. Each resource $n \in \mathcal{N}$ is characterized by upper and lower bounds $\overline{\mathrm{P}}_n, \underline{\mathrm{P}}_n$ over its active power injection, where $\overline{\mathrm{P}}_n, \underline{\mathrm{P}}_n$ can be negative for resources that only consume energy. A resource can be dispatched at any level $x_{n,t}\overline{\mathrm{P}}_n$, such that

$$0 \le x_{n,t}\overline{\mathrm{P}}_n \le \overline{\mathrm{P}}_n, \quad \forall n \in \mathcal{N}, t \in \mathcal{T}, \tag{1}$$

where $x_{n,t} \in [0,1]$ is a decision variable. For a RES facility in particular, the (maximum) power generation at t is a random variable $\widetilde{\mathrm{G}}_{n,t}$, further constraining its dispatch to

$$x_{n,t}\overline{\mathrm{P}}_n \le \widetilde{\mathrm{G}}_{n,t}, \quad \forall n \in \mathcal{N}_{\mathtt{RES}}, t \in \mathcal{T}. \tag{2}$$

Ramp-up and ramp-down constraints for generators read

$$\mathrm{RD}_n \le x_{n,t}\overline{\mathrm{P}}_n - x_{n,t-1}\overline{\mathrm{P}}_{n,t-1} \le \mathrm{RU}_n, \quad \forall n \in \mathcal{N}_{\mathtt{Gen}}, t \in \mathcal{T}. \tag{3}$$

Reactive power injections from consumers and RES are assumed to follow a constant power factor, as in

$$Q_{n,t} = x_{n,t}\overline{\mathrm{P}}_n \tan\left(\cos^{-1}\left(\mathrm{pf}_n\right)\right), \forall n \in \mathcal{N}_{\mathtt{Cons}} \cup \mathcal{N}_{\mathtt{RES}}. \tag{4}$$

On the other hand, for EVs, storage units, and generators, it is assumed that the reactive power can vary within an operational power factor, as in

$$|Q_{n,t}| \le x_{n,t}\overline{\mathrm{P}}_n \tan\left(\cos^{-1}\left(\mathrm{pf}_n\right)\right), \quad \forall n \in \mathcal{N}_{\mathtt{EVs}} \cup \mathcal{N}_{\mathtt{Sto}} \cup \mathcal{N}_{\mathtt{Gen}}, \tag{5}$$

Storage units and EVs have the ability to charge or discharge (both with an assumed efficiency of 1 in this paper), thus it is $\underline{\mathrm{P}}_n \le 0$ and $\overline{\mathrm{P}}_n \ge 0$. A storage unit is characterized by limits $\underline{\mathrm{E}}_n, \overline{\mathrm{E}}_n$ on its battery's energy containment, as in

$$\underline{\mathrm{E}}_n \le \mathrm{E}_{n,0} + \sum_{\tau\in[1,t]} x_{n,\tau}\overline{\mathrm{P}}_n \le \overline{\mathrm{E}}_n, \quad \forall t \in \mathcal{T}, n \in \mathcal{N}_{\mathtt{Sto}}, \tag{6}$$

where $\mathrm{E}_{n,0}$ is the battery's initial energy. We also impose that the battery's state of charge at the end of the horizon is equal to the battery's initial state of charge, by setting

$$\sum_{t\in\mathcal{T}} x_{n,t}\overline{\mathrm{P}}_n = 0, \quad \forall n \in \mathcal{N}_{\mathtt{Sto}}. \tag{7}$$

An EV, on the other hand, is characterized by its arrival and departure times $\widetilde{\mathrm{t}}_n^{\mathrm{arr}}, \widetilde{\mathrm{t}}_n^{\mathrm{dep}}$, its initial state of charge $\widetilde{\mathrm{E}}_n^{\mathrm{arr}}$, its battery capacity $\widetilde{\mathrm{C}}_n$, and a desired state of charge $\widetilde{\mathrm{E}}_n^{\mathrm{des}}$ at its departure time. All of these characteristics are random variables, constituting the EV's type $\mathcal{E}_n$, and are revealed only once the EV arrives in the system. The EV's energy constraint takes the form

$$0 \le \widetilde{\mathrm{E}}_n^{\mathrm{arr}} + \sum_{\tau\in[1,t]} x_{n,\tau}\overline{\mathrm{P}}_n \le \widetilde{\mathrm{C}}_n, \quad \forall t \in \mathcal{T}, n \in \mathcal{N}_{\mathtt{EVs}}, \tag{8}$$

and the difference between the EV's desired state of charge and the actual state of charge at departure is defined as

$$y_n = \widetilde{\mathrm{E}}_n^{\mathrm{des}} - \Big(\widetilde{\mathrm{E}}_n^{\mathrm{arr}} + \sum_{t\in[\widetilde{\mathrm{t}}_n^{\mathrm{arr}}, \widetilde{\mathrm{t}}_n^{\mathrm{dep}}]} x_{n,t}\overline{\mathrm{P}}_n\Big), \quad \forall n \in \mathcal{N}_{\mathtt{EVs}}. \tag{9}$$

Moreover, for exactness, we write:

$$x_{n,t} = 0, \quad \forall n \in \mathcal{N}_{\mathtt{EVs}}, t \notin [\widetilde{\mathrm{t}}_n^{\mathrm{arr}}, \widetilde{\mathrm{t}}_n^{\mathrm{dep}}]. \tag{10}$$

Each resource can control its dispatch profile $\boldsymbol{x}_n \triangleq (x_{n,t})_{t\in\mathcal{T}}$ over the horizon $\mathcal{T}$, at a cost given by the resource's cost function $c_n(\boldsymbol{x}_n)$. In this paper, we model the resources' cost functions as convex functions of $\boldsymbol{x}_n$. In particular, the cost functions for generators take the form

$$c_{n:n\in\mathcal{N}_{\mathtt{Gen}}}(\boldsymbol{x}_n) = \sum_{t\in\mathcal{T}} \mathrm{g}_n \cdot (x_{n,t}\overline{\mathrm{P}}_n)^2 - \widetilde{\mathrm{p}}_t \cdot x_{n,t}\overline{\mathrm{P}}_n, \tag{11}$$

where g_n relates to the fuel cost, and $\widetilde{\mathrm{p}}_t$ is the wholesale market price (a random variable), yielding generation revenues.

A RES facility, gains wholesale market revenues with virtually zero operational cost. Thus, its cost function is decreasing in $x_{n,t}$, as in

$$c_{n:n\in\mathcal{N}_{\mathtt{RES}}}(\boldsymbol{x}_n) = -\sum_{t\in\mathcal{T}} (\widetilde{\mathrm{p}}_t) \cdot x_{n,t}\overline{\mathrm{P}}_n. \tag{12}$$

Storage units are also subject to wholesale market revenues (or payments for $x_{n,t}\overline{\mathrm{P}}_n < 0$), and additionally bear a battery degradation cost, as in

$$c_{n:n\in\mathcal{N}_{\mathtt{Sto}}}(\boldsymbol{x}_n) = \sum_{t\in\mathcal{T}} \left(\mathrm{d} \cdot \left(\frac{x_{n,t}\overline{\mathrm{P}}_n}{\overline{\mathrm{E}}_n}\right)^2 - \widetilde{\mathrm{p}}_t \cdot x_{n,t}\overline{\mathrm{P}}_n\right), \tag{13}$$

where d is a battery degradation factor, and $\frac{\sum_{t\in\mathcal{T}} x_{n,t}\overline{\mathrm{P}}_n}{\overline{\mathrm{E}}_n}$ is the number of full charge–discharge cycles.

EVs bear the same costs as storage units, and an additional penalty/disutility cost $u(y_n)$ (e.g. quadratic in y_n) for not having their battery charged at their desired level upon departure:

$$c_{n:n\in\mathcal{N}_{\mathtt{EVs}}}(\boldsymbol{x}_n) = \sum_{t\in\mathcal{T}} \left(\mathrm{d} \cdot \left(\frac{x_{n,t}\overline{\mathrm{P}}_n}{\widetilde{\mathrm{C}}_n}\right)^2 - \widetilde{\mathrm{p}}_t \cdot x_{n,t}\overline{\mathrm{P}}_n\right) + u(y_n). \tag{14}$$

Finally, a consumer has an energy demand level $\widetilde{\mathrm{D}}_{n,t}$ at each timeslot. In addition to its retail cost $\sum_{t\in\mathcal{T}} \widetilde{\mathrm{p}}_t \cdot x_{n,t}\overline{\mathrm{P}}_n$, it bears an instantaneous cost $\mathrm{w}_2 \cdot \left(\widetilde{\mathrm{D}}_{n,t} - x_{n,t}\overline{\mathrm{P}}_n\right)^2$ for having its load shifted from (or to) timeslot t, as well as an extra cost $\mathrm{w}_3 \cdot \left(\sum_{t\in\mathcal{T}} \widetilde{\mathrm{D}}_{n,t} - \sum_{t\in\mathcal{T}} x_{n,t}\overline{\mathrm{P}}_n\right)^2$ for having part of its demand unsatisfied:

$$c_{n:n\in\mathcal{N}_{\mathtt{Cons}}}(\boldsymbol{x}_n) = \sum_{t\in\mathcal{T}} \left(\widetilde{\mathrm{p}}_t \cdot x_{n,t}\overline{\mathrm{P}}_n + \mathrm{w}_2 \cdot (\widetilde{\mathrm{D}}_{n,t} - x_{n,t}\overline{\mathrm{P}}_n)^2\right) + \mathrm{w}_3 \cdot \left(\sum_{t\in\mathcal{T}} \widetilde{\mathrm{D}}_{n,t} - \sum_{t\in\mathcal{T}} x_{n,t}\overline{\mathrm{P}}_n\right)^2. \tag{15}$$

2.2. Distribution system

The resources are connected via a radial distribution network defined by the set of nodes/buses $\mathcal{B}$ and their interconnecting lines. For a bus $b \in \mathcal{B}$, we denote the set of resources connected to it by $\mathcal{N}_b$, its parent node by ζ_b and the set of its children nodes by $\mathcal{C}_b$. The active and reactive power flows from node i to node j, at t, are denoted as $P_{ij,t}$ and $Q_{ij,t}$ respectively. A node's active power balance is ensured by

$$P_{\zeta_b b,t} + \sum_{n\in\mathcal{N}_b} x_{n,t}\overline{\mathrm{P}}_n - \sum_{c\in\mathcal{C}_b} \left(P_{bc,t} + \mathrm{R}_{bc} I_{bc,t}^{\mathrm{sqr}}\right) = 0 \quad \forall\, b \in \mathcal{B},\, t \in \mathcal{T}, \tag{16}$$

where $I_{bc,t}^{\mathrm{sqr}}$ is the squared magnitude of the current flowing through the line connecting b to c, and R_{bc} is the line's resistance. Similarly, the reactive power balance is written as

$$Q_{\zeta_b b,t} + \sum_{n\in\mathcal{N}_b} Q_{n,t} - \sum_{c\in\mathcal{C}_b} \left(Q_{bc,t} + \mathrm{X}_{bc} I_{bc,t}^{\mathrm{sqr}}\right) = 0 \quad \forall\, b \in \mathcal{B},\, t \in \mathcal{T}, \tag{17}$$

where X_{bc} is the line's reactance. The voltage magnitude drop between nodes ζ_b and b is represented by:

$$V_{\zeta_b,t}^{\mathrm{sqr}} - 2\left(\mathrm{R}_{\zeta_b b} P_{\zeta_b b,t} + \mathrm{X}_{\zeta_b b} Q_{\zeta_b b,t}\right) - \left(\mathrm{R}_{\zeta_b b}^2 + \mathrm{X}_{\zeta_b b}^2\right) I_{\zeta_b b,t}^{\mathrm{sqr}}$$

$$= V^{\rm sqr}_{b,t}, \quad \forall b \in \mathcal{B}, t \in \mathcal{T}, \tag{18}$$

while branch power flows are calculated using the (SOCP-relaxed) inequality

$$V_{b,t} I_{\zeta_b b,t} \leq P^2_{\zeta_b b,t} + Q^2_{\zeta_b b,t}, \qquad \forall b \in \mathcal{B}, t \in \mathcal{T}. \tag{19}$$

Finally, the upper and lower bounds on nodal voltage magnitudes and current magnitudes are enforced by

$$\underline{\rm V} \leq V_{b,t} \leq \overline{\rm V} \qquad \forall b \in \mathcal{B}, t \in \mathcal{T}, \tag{20}$$

$$0 \leq I_{\zeta_b b,t} \leq \bar{\rm I}_{\zeta_b b} \qquad \forall b \in \mathcal{B}, t \in \mathcal{T}. \tag{21}$$

2.3. Problem formulation

Under no uncertainties over the system's parameters (e.g. demand and RES output), the optimal-in-hindsight solution is given by the following OPF problem:

$$\begin{aligned} &\min_{\mathcal{U}} \sum_{n \in \mathcal{N}} c_n(\boldsymbol{x}_n) \\ &\text{s.t. (1)–(21),} \end{aligned} \tag{22}$$

where the set of decision variables is

$$\mathcal{U} = \{ (x_{n,t}, Q_{n,t})_{n \in \mathcal{N}, t \in \mathcal{T}}, (y_n)_{n \in \mathcal{N}}, (V_{b,t}, P_{\zeta_b b,t}, Q_{\zeta_b b,t}, I_{\zeta_b b,t})_{b \in \mathcal{B}, t \in \mathcal{T}} \}.$$

However, the problem's parameters include a set

$$\mathcal{W}_t = \{ (\widetilde{\rm G}_{n,t})_{n \in \mathcal{N}_{\tt RES}}, (\widetilde{\rm D}_{n,t})_{n \in \mathcal{N}_{\tt Cons}}, \widetilde{\rm p}_t \} \cup (\mathcal{E}_n)_{n \in \mathcal{N}_{\tt EVs} : \widetilde{\rm t}^{\rm arr}_n = t} \tag{23}$$

of random variables for each t (i.e. the RES generation, consumer demand, prices, and EV characteristics) rendering the tracking of the optimal dispatch a problem of sequential decision making under uncertainty. Using the general modeling framework for such problems, as introduced in [13], our problem is defined by:

- The set of decision stages $\mathcal{T}$.
- The set of action variables at each stage:

  $$\mathcal{U}_t = \{ (x_{n,t}, Q_{n,t})_{n \in \mathcal{N}}, (V_{b,t}, P_{\zeta_b b,t}, Q_{\zeta_b b,t}, I_{\zeta_b b,t})_{b \in \mathcal{B}} \}.$$

- The system's *state* at t:

  $$\mathcal{S}_t = \{ (x_{n,t-1})_{n \in \mathcal{N}_{\tt Gen}}, (E_{n,t-1})_{n \in \mathcal{N}_{\tt EVs} \cup \mathcal{N}_{\tt Sto} \cup \mathcal{N}_{\tt Cons}}, \mathcal{W}_t, \mathcal{F}_t \}$$

  which represents all the information relevant for making a decision; this includes each generator's previous output level $x_{n,t-1}$, a state-of-energy variable

  $$E_{n,t-1} = \sum_{t'=1}^{t-1} x_{n,t'} \tag{24}$$

  for storage, EVs and consumers, the currently revealed information $\mathcal{W}_t$, and a belief-state $\mathcal{F}_t$ which encompasses all the parameters relevant for reasoning over the future realizations $(\mathcal{W}_{t'})_{t' \in [t+1,|\mathcal{T}|]}$ of the system's random variables.
- The system's stage cost $C_t(\mathcal{U}_t, \mathcal{S}_t)$ defined as the sum of the resources' cost functions.
- A transition function H that maps $(\mathcal{U}_t, \mathcal{S}_t)$ to a next state $\mathcal{S}_{t+1}$; this comprises the deterministic transition functions of state components $x_{n,t-1}, E_{n,t-1}$, the unknown dynamics of the random variables $\mathcal{W}_t$, and the method-specific dynamics that define the update rules for the parameters of the belief-state $\mathcal{F}_t$.

The solution concept for our problem takes the form of a *policy* π, i.e. a method $\mathcal{U}_t = \pi(\mathcal{S}_t)$ for deciding feasible actions $\mathcal{U}_t$ at any realization of the state $\mathcal{S}_t$. Based on these definitions, our objective can be defined as the minimization (over policies) of the expected system's accumulated cost:

$$\begin{aligned} &\min_{\pi} \left\{ \sum_{t \in \mathcal{T}} \mathbb{E}_{\psi \sim \pi} \left[C_t(\mathcal{U}^{\pi}_t, \mathcal{S}_t) \right] \right\} \\ &\text{s.t.} \quad \mathcal{S}_{t+1} = H^{\pi}(\mathcal{U}^{\pi}_t, \mathcal{S}_t) \end{aligned} \tag{25}$$

where the expectation is over the system's possible state–action trajectories ψ conditioned on the adopted policy π.

In the next section, we present two benchmark policies for problem (25): a receding horizon stochastic program and a learn-to-optimize approach. These policies serve as building blocks for the proposed policy (to be presented later in Section 4) and also as benchmarks against which the proposed policy will be evaluated in Section 5.

3. Benchmark policies

This section presents two benchmark policies for problem (25) before presenting the proposed policy in the next Section.

3.1. Receding-horizon stochastic programming

Let us consider a given decision stage and denote it by τ. Given the revealed information $\mathcal{W}_\tau$, a stochastic programming approach considers a set $\mathcal{K}$ of *scenarios* for future realizations $(\mathcal{W}_{t,k})_{t \in [\tau+1,|\mathcal{T}|]}$ of the system's random variables. These scenarios constitute the policy's belief-state $\mathcal{F}_\tau$ at τ and are generated by drawing on the statistical properties of past observations. The program uses a duplicate set $\mathcal{U}_{t,k}$ of decision variables for each scenario and timeslot, and makes a decision $\mathcal{U}^{\rm SP}_\tau$ at τ by solving the following optimization problem:

$$\min_{(\mathcal{U}_{t,k})_{t \in \mathcal{T}, k \in \mathcal{K}}} \quad \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} c_n(\boldsymbol{x}_{n,k}) \tag{26}$$

$$\text{s.t.} \quad (1)\text{–}(21), \quad \forall k \in \mathcal{K} \tag{27}$$

$$(\mathcal{U}_{t,k})_{t \in [1,\tau-1], k \in \mathcal{K}} = (\mathcal{U}^{\rm SP}_t)_{t \in [1,\tau-1]}, \tag{28}$$

$$(\mathcal{U}_{\tau,k})_{k \in \mathcal{K}} = (\mathcal{U}_{\tau,k'})_{k' \in \mathcal{K}}, \quad \forall k, k' \in \mathcal{K}, \tag{29}$$

where (27) enforces the system's operational constraints for each scenario, constraint (28) fixes the decisions made before current stage t to the applied actions (ensuring that the algorithm cannot change the past), and constraint (29) enforces the so-called non-anticipativity constraints. At stage τ, and given an optimal solution to the above problem, the receding-horizon algorithm applies the decision for the current stage and re-solves the optimization in the next stage $\tau + 1$ when the belief-state is updated by considering the newly revealed information $\mathcal{W}_{\tau+1}$.

3.2. Learn to optimize

The learn-to-optimize approach avoids the need to solve a stochastic OPF problem in online operation, by feeding the observed state to a ML algorithm, namely a neural network (NN), which provides an estimation of the optimal dispatch in negligible time. More specifically, at stage τ of online operation, the NN is provided with the input

$$\mathcal{I}_\tau = \{ \tau, (x_{n,\tau-1})_{n \in \mathcal{N}_{\tt Gen}}, (E_{n,\tau-1})_{n \in \mathcal{N}_{\tt EVs} \cup \mathcal{N}_{\tt Sto} \cup \mathcal{N}_{\tt Cons}}, (\mathcal{W}_t)_{t \in [1,\tau]} \}, \tag{30}$$

i.e., all the relevant information currently available, and provides as output an estimation $\mathcal{O}_\tau = (x^{\tt LtO}_{n,\tau}, Q^{\tt LtO}_{n,\tau})_{n \in \mathcal{N}}$ of the optimal dispatch (where the LtO superscript specifies that this the solution prescribed by the Learn-to-Optimize approach. The NN is trained offline, using mappings of the form

$$m^d_t = \left(\mathcal{I}^d_t, \mathcal{O}^d_t \right). \tag{31}$$

To create a mapping m^d_t, an instance $\mathcal{W}^d = \left(\mathcal{W}^d_t \right)_{t \in \mathcal{T}}$ for a whole day is considered and the respective optimal-in-hindsight solution $\mathcal{U}^d$ is calculated by solving problem (22) (under a perfect forecast). Thus, for each t of day d, the input feature $x^d_{n,t-1}$ is the "$(n, t-1)$" component of the optimal solution $\mathcal{U}^d$ and $E^d_{n,t-1}$ is calculated using (24). The output part $\mathcal{O}^d_t$ of m^d_t is simply the optimal value of variables $x_{n,t}, Q_{n,t} \in \mathcal{U}^d$.

One shortcoming of this policy is that, for a large number $|\mathcal{N}|$ of resources, the dimension of the NN's output also grows, which obstructs the NN's efficient training and performance. A second shortcoming refers to the inability of the policy to guarantee the feasibility of the control actions. The next Section presents the proposed policy which remedies these issues.

4. Proposed policy

In this Section, we present the proposed distributed policy for problem (25). Let us first consider a distributed algorithm for the perfect forecast case (refer to problem (22)) before we address uncertainty. Let us consider the active and reactive power balance residuals, for each node and timeslot, as

$$\delta^{\mathrm{p}}_{b,t} = P_{\zeta_b b,t} + \sum_{n\in\mathcal{N}_b} x_{n,t}\overline{\mathrm{P}}_n - \sum_{c\in\mathcal{C}_b}\left(P_{bc,t} + \mathrm{R}_{bc} I^{\mathrm{sqr}}_{bc,t}\right), \tag{32}$$

$$\delta^{\mathrm{q}}_{b,t} = Q_{\zeta_b b,t} + \sum_{n\in\mathcal{N}_b} Q_{n,t} - \sum_{c\in\mathcal{C}_b}\left(Q_{bc,t} + \mathrm{X}_{bc} I^{\mathrm{sqr}}_{bc,t}\right). \tag{33}$$

By relaxing the respective active and reactive power balance constraints (16), (17), we can write the augmented Lagrangian of problem (22) as:

$$L = \sum_{n\in\mathcal{N}} c_n(\boldsymbol{x}_n) - \sum_{t\in\mathcal{T}}\sum_{b\in\mathcal{B}}\left(\lambda_{b,t}\cdot\delta^{\mathrm{p}}_{b,t} - \frac{\rho}{2}\left(\delta^{\mathrm{p}}_{b,t}\right)^2\right) - \sum_{t\in\mathcal{T}}\sum_{b\in\mathcal{B}}\left(\mu_{b,t}\cdot\delta^{\mathrm{q}}_{b,t} - \frac{\rho}{2}\left(\delta^{\mathrm{q}}_{b,t}\right)^2\right), \tag{34}$$

where $\lambda_{b,t}, \mu_{b,t}$ are the Lagrange multipliers of the active and reactive power balance constraints for b, t. Given that the Lagrangian is per-node separable, the optimal-in-hindsight problem (22) lends itself to a distributed solution where, at iteration i, each node updates its decisions $\mathcal{U}_b = \left((x_{n,t}, Q_{n,t})_{n\in\mathcal{N}_b, t\in\mathcal{T}}, (y_n)_{n\in\mathcal{N}_b}\right)$ as

$$\mathcal{U}^{(i)}_b \in \mathbf{argmin}_{\mathcal{U}}\ \{L\} \quad \text{s.t. } (1)\text{–}(15), \quad \mathcal{U}_{\mathrm{DSO}} = \mathcal{U}^{(i-1)}_{\mathrm{DSO}} \tag{35}$$

and the DSO updates its variables

$$\mathcal{U}_{\mathrm{DSO}} = \left(V_{b,t}, P_{\zeta_b b,t}, Q_{\zeta_b b,t}, I_{\zeta_b b,t}\right)_{b\in\mathcal{B}, t\in\mathcal{T}}$$

as

$$\mathcal{U}^{(i)}_{\mathrm{DSO}} \in \mathbf{argmin}_{\mathcal{U}}\ \{L\} \quad \text{s.t. } (18)\text{–}(21), \quad \mathcal{U}_b = \mathcal{U}^{(i-1)}_b, \quad \forall b\in\mathcal{B}. \tag{36}$$

Given the simultaneous variables' updates (35), (36), the Lagrange multipliers can be updated using the Alternate Direction Method of Multipliers (ADMM), as:

$$\lambda^{(i)}_{b,t} = \lambda^{(i-1)}_{b,t} + \rho\delta^{\mathrm{p}}_{b,t}\left(\mathcal{U}^{(i)}_b, \mathcal{U}^{(i)}_{\mathrm{DSO}}\right), \qquad \forall b\in\mathcal{B}, t\in\mathcal{T}, \tag{37}$$

$$\mu^{(i)}_{b,t} = \mu^{(i-1)}_{b,t} + \rho\delta^{\mathrm{q}}_{b,t}\left(\mathcal{U}^{(i)}_b, \mathcal{U}^{(i)}_{\mathrm{DSO}}\right) \qquad \forall b\in\mathcal{B}, t\in\mathcal{T}. \tag{38}$$

We now turn to constructing the proposed policy for the sequential decision-making problem under uncertainty. Similarly to the learn-to-optimize policy of Section 3.2, the proposed policy is assisted by a NN trained on instances of the problem's optimal-in-hindsight solution. However, instead of training the NN to estimate the optimal dispatch at current stage τ, it is trained to return an estimation $(\lambda^*_{b,t}, \mu^*_{b,t})_{t\in[\tau+1,|\mathcal{T}|]}$ of the problem's optimal *dual* variables that correspond to constraints (16), (17) of problem (22). Thus, at stage τ of online operation, the NN is fed with the same input $\mathcal{I}_\tau$ as defined in (30), plus the dual variables in which the ADMM algorithm converged in the previous timeslot, and predicts the optimal dual variables for future stages. Then, the ADMM algorithm is executed, where only the multipliers for the current timeslot τ are iteratively updated, while the multipliers for future timeslots are kept fixed to the estimated values $(\lambda^*_{b,t}, \mu^*_{b,t})_{t\in[\tau+1,|\mathcal{T}|]}$. The exact policy at decision stage τ, reads as in Algorithm 1.

Algorithm 1 The proposed policy for the data-driven, distributed sequential OPF problem at stage τ.

Feed the input (Eq. (30)) to the NN, and obtain the predicted multipliers $\lambda^*_{b,t}, \mu^*_{b,t}$, for each node $b \in \mathcal{B}$, and future stage $t \in [\tau+1, |\mathcal{T}|]$
Fix the multipliers for future stages to the ones predicted by the NN: $\lambda_{b,t}, \mu_{b,t} = \lambda^*_{b,t}, \mu^*_{b,t}, \quad \forall b\in\mathcal{B}, t\in[\tau+1,|\mathcal{T}|]$.
Initialize the iteration number and the multipliers for the current stage $i = 0$, $\lambda^{(0)}_{b,\tau}, \mu^{(0)}_{b,\tau} = 0, \quad \forall b\in\mathcal{B}$.
repeat:
  $i = i + 1$
  for $b \in \mathcal{B}$:
    Update node's decisions $\mathcal{U}^{(i)}_b$ by solving (35)
  Update DSO decisions $\mathcal{U}^{(i)}_{\mathrm{DSO}}$ by solving (36)
  Update the multipliers for the current stage, as:
  $\lambda^{(i)}_{b,\tau} = \lambda^{(i-1)}_{b,\tau} + \rho\delta^{\mathrm{p}}_{b,\tau}\left(\mathcal{U}^{(i)}_b, \mathcal{U}^{(i)}_{\mathrm{DSO}}\right), \quad \forall b\in\mathcal{B}$,
  $\mu^{(i)}_{b,\tau} = \mu^{(i-1)}_{b,\tau} + \rho\delta^{\mathrm{q}}_{b,\tau}\left(\mathcal{U}^{(i)}_b, \mathcal{U}^{(i)}_{\mathrm{DSO}}\right) \quad \forall b\in\mathcal{B}$.
until $\max_{b\in\mathcal{B}}\{\lambda^{(i)}_{b,\tau} - \lambda^{(i-1)}_{b,\tau}\} < \epsilon$
  AND
  $\max_{b\in\mathcal{B}}\{\mu^{(i)}_{b,\tau} - \mu^{(i-1)}_{b,\tau}\} < \epsilon$
apply: $(x^{(i)}_{n,\tau}, Q^{(i)}_{n,\tau})_{n\in\mathcal{N}_b}$

Remark 1. Notice that the proposed policy of Algorithm 1 converges to an uncertainty-informed dispatch that always respects the system's constraints (by construction).

5. Experimental evaluation

5.1. Evaluation setup

The presented policies are evaluated for a 24-timeslots horizon on the 11 kV MV-distribution network of [14]. The substation voltage was set to 1.0 p.u. on the secondary side and voltage magnitude limits to $\overline{\mathrm{V}} = 1.05$ p.u. and $\underline{\mathrm{V}} = 0.95$ p.u.

The dataset from [15] was used for the consumers' consumption. For EVs, we assumed that each user wants to charge his/her EV as much as possible within the given deadline and the disutility cost is linear in the amount of energy not charged. The dataset from [16] was used for EVs' characteristics. Wholesale market prices were drawn from [17], and RES generation was drawn from [18]. For the proposed policy, a Graph Neural Network was used to predict the dual variables. The NN consists of two graph convolution layers and 3 linear layers. The layers use the Relu activation function. Each bus/node n is represented with its local state variable which contains only the variables that refer to that bus/node.

5.2. Evaluation results

Our main result is the comparison of the proposed policy of Algorithm 1 against the optimal in hindsight solution (*oracle*) of problem (22) and against the two benchmark policies (the receding-horizon stochastic program and the learn-to-optimize policy). Naturally, the optimal-in-hindsight solution provides us with the ideal objective value that could only be reached if all information was known beforehand

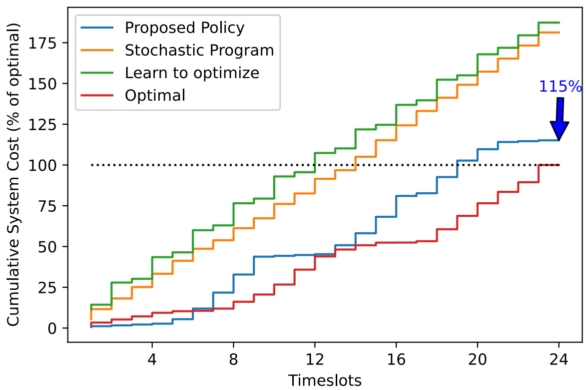

Fig. 1. Comparison of the three policies with the optimal-in-hindsight solution.

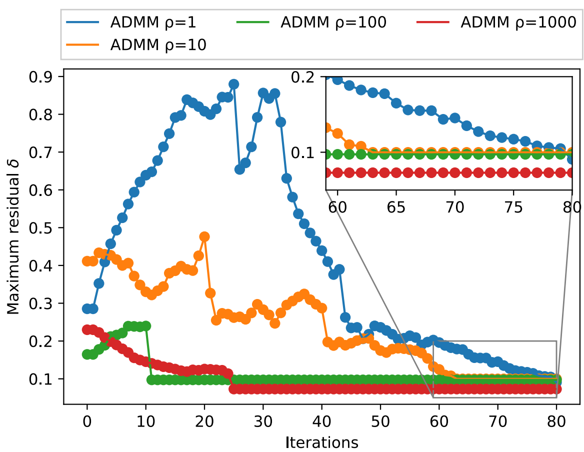

Fig. 2. Convergence behavior of the proposed policy.

and it is not attainable in practice. Nevertheless, it serves as a theoretical benchmark against which we can assess the performance of the three policies. In Fig. 1, we present each policy's system cost accumulated along the horizon, where the value at the last timeslot (t = 24) expresses the policy's overall performance as a percentage of the optimal-in-hindsight solution's cost. As can be observed, the proposed policy significantly outperforms the two benchmarks by achieving a cost that is only 15% higher than the one of the perfect information case.

The performance documented in Fig. 1 for the proposed policy was achieved by setting ρ equal to 1. Higher values of ρ can provide faster convergence times as can be seen in Fig. 2, although at the expense of higher system cost (loss of efficiency) which can be significant as shown in Fig. 3. However, for the near-optimal choice of $\rho = 1$, the computational time required to make a decision was in the order of only one minute, which is already fast enough for the intended application and validates the policy's suitability for real-time decisions. Further simulation results, including e.g. a sensitivity analysis to biased data are included in the extended online version of the paper [19].

6. Conclusion and future work

This paper motivated the need to solve the distribution-level OPF problem in a stochastic, sequential, distributed and data-driven manner. The problem was formulated as a problem of sequential decision-making under uncertainty and the notion of a policy was identified as the relevant solution concept. A policy was proposed which uses a Graph Neural Network to predict the problem's optimal dual variables for future intervals combined with an iterative distributed optimization

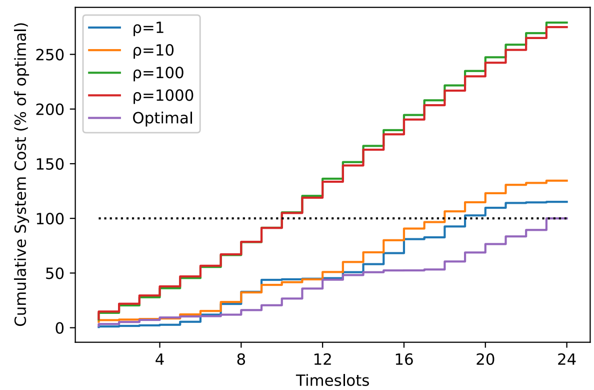

Fig. 3. Efficiency loss for different values of parameter ρ.

algorithm for making coordinated real-time decisions that, by construction, respect the system's operational constraints. The proposed policy was compared to two policies commonly used in the literature (a stochastic programming and a direct machine-learning approach) demonstrating a significant difference in the achieved efficiency. The simulation results indicated the ability of the proposed policy to make fast online decisions while approaching the objective value of the optimal (perfect information) solution. The proposed methodology is not confined by the distribution network's (radial) structure and can also work for meshed networks in principle. However, how well it would performed in meshed networks is an open question that can be empirically evaluated in future work.

CRediT authorship contribution statement

Georgios Tsaousoglou: Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing. **Petros Ellinas:** Data curation, Investigation, Software, Validation, Visualization. **Juan S. Giraldo:** Methodology, Software, Validation, Writing – review & editing. **Emmanouel Varvarigos:** 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.

References

[1] S. Chatzivasileiadis, A. Venzke, J. Stiasny, G. Misyris, Machine learning in power systems: Is it time to trust it? IEEE Power Energy Mag. 20 (3) (2022) 32–41.
[2] D.K. Molzahn, F. Dörfler, H. Sandberg, S.H. Low, S. Chakrabarti, R. Baldick, J. Lavaei, A survey of distributed optimization and control algorithms for electric power systems, IEEE Trans. Smart Grid 8 (6) (2017) 2941–2962.
[3] G. Tsaousoglou, P. Soumplis, N. Efthymiopoulos, K. Steriotis, A. Kretsis, P. Makris, P. Kokkinos, E. Varvarigos, Demand response as a service: Clearing multiple distribution-level markets, IEEE Trans. Cloud Comput. 10 (1) (2021) 82–96.
[4] A. Rajaei, S. Fattaheian-Dehkordi, M. Fotuhi-Firuzabad, M. Moeini-Aghtaie, M. Lehtonen, Developing a distributed robust energy management framework for active distribution systems, IEEE Trans. Sustain. Energy 12 (4) (2021) 1891–1902.
[5] Y. Du, W. Pei, N. Chen, X. Ge, H. Xiao, Real-time microgrid economic dispatch based on model predictive control strategy, J. Mod. Power Syst. Clean Energy 5 (5) (2017) 787–796.
[6] M. Bazrafshan, N. Gatsis, Decentralized stochastic optimal power flow in radial networks with distributed generation, IEEE Trans. Smart Grid 8 (2) (2016) 787–801.

[7] C. Wang, S. Lei, P. Ju, C. Chen, C. Peng, Y. Hou, MDP-based distribution network reconfiguration with renewable distributed generation: Approximate dynamic programming approach, IEEE Trans. Smart Grid 11 (4) (2020) 3620–3631.

[8] M. Chertkov, D. Deka, Y. Dvorkin, Optimal ensemble control of loads in distribution grids with network constraints, in: 2018 Power Systems Computation Conference, PSCC, IEEE, 2018, pp. 1–7.

[9] X. Pan, M. Chen, T. Zhao, S.H. Low, DeepOPF: A feasibility-optimized deep neural network approach for AC optimal power flow problems, IEEE Syst. J. (2022).

[10] G. Chen, H. Zhang, H. Hui, N. Dai, Y. Song, Scheduling thermostatically controlled loads to provide regulation capacity based on a learning-based optimal power flow model, IEEE Trans. Sustain. Energy 12 (4) (2021) 2459–2470.

[11] X. Lei, Z. Yang, J. Yu, J. Zhao, Q. Gao, H. Yu, Data-driven optimal power flow: A physics-informed machine learning approach, IEEE Trans. Power Syst. 36 (1) (2020) 346–354.

[12] M.K. Singh, V. Kekatos, G.B. Giannakis, Learning to solve the AC-OPF using sensitivity-informed deep neural networks, IEEE Trans. Power Syst. 37 (4) (2021) 2833–2846.

[13] W.B. Powell, A unified framework for stochastic optimization, European J. Oper. Res. 275 (3) (2019) 795–821.

[14] LaPSEE Power System Test Cases Repository UNESP, Downloads / sistemas testes, 2020, https://bit.ly/2IdSn91. (Online; Accessed October 2020).

[15] P. Staudt, N. Ludwig, J. Huber, V. Hagenmeyer, C. Weinhardt, Sciber: A new public data set of municipal building consumption, in: Proceedings of the Ninth International Conference on Future Energy Systems, 2018, pp. 618–621.

[16] Adaptive charging network, 2021, https://ev.caltech.edu/dataset. (Accessed 10 2021).

[17] Results of the Greek balancing market, 2021, https://www.admie.gr/en/market/market-statistics/detail-data. (Accessed 10 2021).

[18] VIMSEN project repository, 2021, https://github.com/vimsen/dss/tree/master/data_points. (Accessed 10 2021).

[19] G. Tsaousoglou, P. Ellinas, J. Giraldo, E. Varvarigos, Distributed sequential optimal power flow under uncertainty in power distribution systems: A data-driven approach, 2022, TechRxiv.