Learning the price response of active distribution networks for TSO-DSO coordination

J. M. Morales, S. Pineda and Y. Dvorkin

Abstract—The increase in distributed energy resources and flexible electricity consumers has turned TSO-DSO coordination strategies into a challenging problem. Existing decomposition/decentralized methods apply divide-and-conquer strategies to trim down the computational burden of this complex problem, but rely on access to proprietary information or fail-safe real-time communication infrastructures. To overcome these drawbacks, we propose in this paper a TSO-DSO coordination strategy that only needs a series of observations of the nodal price and the power intake at the substations connecting the transmission and distribution networks. Using this information, we learn the price response of active distribution networks (DN) using a decreasing step-wise function that can also adapt to some contextual information. The learning task can be carried out in a computationally efficient manner and the curve it produces can be interpreted as a market bid, thus averting the need to revise the current operational procedures for the transmission network. Inaccuracies derived from the learning task may lead to suboptimal decisions. However, results from a realistic case study show that the proposed methodology yields operating decisions very close to those obtained by a fully centralized coordination of transmission and distribution.

Index Terms—TSO-DSOs coordination, DERs market integration, distribution network, price-responsive consumers, statistical learning.

Nomenclature

The main symbols used throughout this paper are listed next for quick reference. Others are defined as required in the text. We consider hourly time steps and, hence, MW and MWh are interchangeable under this premise.

A. Indexes and sets

b	Index of blocks in the step-wise approximation of the price response function.
i	Index of generating units.
j	Index of consumers.
k	Index of distribution networks.
l	Index of lines.

J. M. Morales is with the Department of Applied Mathematics, University of Málaga, Málaga, Spain. E-mail: juan.morales@uma.es

S. Pineda is with the Department of Electrical Engineering, University of Málaga, Málaga, Spain. E-mail: spinedamorente@gmail.com.

Y. Dvorkin is with the New York University, Brooklyn, NY 11201 USA. E-mail: dvorkin@nyu.edu.

n	Index of nodes.
t	Index of time periods.
$\mathcal{B}$	Set of blocks in the step-wise approximation of the price response function.
I^T	Set of generating units at the transmission network.
I_n	Set of generating units at node n.
I_k^D	Set of generating units of distribution network k.
J^T	Set of consumers in the transmission network.
J_n	Set of consumers at node n.
J_k^D	Set of consumers in distribution network k.
K^T	Set of distribution networks.
K_n	Set of distribution networks at node n.
L^T	Set of lines of the transmission network.
L_k^D	Set of lines of distribution network k.
N^T	Set of nodes of the transmission network.
N_k^D	Set of nodes of distribution network k.
$\mathcal{T}$	Set of time periods.
$\mathcal{T}^C(t)$	Subset of time periods that are the closest to t.

B. Parameters

a_i	Quadratic cost parameter of unit i [€/MW^2].
b_i	Linear cost parameter of unit i [€/MW].
$\underline{p}_i^G$	Minimum active power output of unit i [MW].
$\overline{p}_i^G$	Maximum active power output of unit i [MW].
$\widehat{p}_{jt}^D$	Baseline demand of consumer j at time t [MW].
$\overline{p}_{jt}^D$	Maximum demand of consumer j at time t [MW].
$\underline{p}_{jt}^D$	Minimum demand of consumer j at time t [MW].
$\underline{q}_i^G$	Minimum reactive power output of unit i [MW].
$\overline{q}_i^G$	Maximum reactive power output of unit i [MW].
r_l	Resistance of line l [p.u.].
$\overline{s}_l^F$	Capacity of line l [MVA].
$\overline{s}_i^G$	Inverter power rate of unit i [MVA].
s^B	Base power [MVA].
$\overline{u}_{ktb}^B$	Marginal utility of block b for DN k and time t [€/MW].
$\overline{v}_n$	Maximum squared voltage at node n [p.u.].
$\underline{v}_n$	Minimum squared voltage at node n [p.u.].
x_l	Reactance of line l [p.u.].
α_{jt}	Intercept of the inverse demand function of consumer j at time t [MW].
β_{jt}	Slope of the inverse demand function of consumer j at time t [€/MW^2].
γ_j	Power factor of consumer j.
δ_j	Flexibility parameter of consumer j.
λ_{kt}	Price at the substation of DN k at time t [€/MW].
ρ_{it}	Capacity factor of unit i at time t.
χ_{kt}	Contextual information of DN k at time t.

C. Variables

p_{it}^G	Active power output of unit i at time t [MW].
p_{jt}^D	Active power demand of consumer j at time t [MW].
p_{lt}^F	Active power flow through line l at time t [MW].
p_{kt}^N	Active power intake of DN k at time t [MW].
q_{it}^G	Reactive power output of unit i at time t [MVAr].
q_{jt}^D	Reactive power demand of consumer j at time t [MVAr].
q_{lt}^F	Reactive power flow through line l at time t [MVAr].
v_{nt}	Squared voltage magnitude at node n and time t [p.u.].
θ_{nt}	Voltage angle at node n and time t [rad].

I. INTRODUCTION

ELECTRIC power distribution has been traditionally ignored in the operation of transmission power networks, on the grounds that distribution grids only housed passive loads. However, the proliferation of distributed energy resources (DERs) is rendering this traditional *modus operandi* obsolete [1]. Power systems engineers are faced with an unprecedented challenge of efficiently integrating a vast number and a wide spectrum of flexible power assets located in mid- and low-voltage networks into the operation of the transmission power network [2]. Naturally, succeeding in this endeavor requires the coordination between the transmission and distribution system operators (TSOs and DSOs, respectively), all united in the purpose of fostering an active role of DERs in the operation of the power system through their participation in wholesale electricity markets.

As a result, research emphasis is placed on mechanisms that strengthen the TSO-DSO coordination so that the available flexibility of DERs can be harvested for transmission and wholesale market services [3], [4]. For instance, some recent research works investigate TSO-DSO coordination schemes to improve voltage stability [5], [6]. Other authors focus on the economic coordination between transmission and distribution operations to minimize total system costs [7]. The present work belongs to this latter group.

Regarding TSO-DSO economic coordination, a single centralized operational model that includes both transmission and distribution networks with their full level of detail is not viable due to its computational cost, modeling complexity and potential conflict of interests between the involved parties [8]. Rather, the coordination of transmission and distribution power assets calls for a divide-and-conquer strategy that alleviates the computational burden, allows for decentralization and minimizes the need for information exchange between the TSO and DSOs [9]. For instance, authors of [10] uses Benders decomposition to find the optimal economic dispatch considering TSO-DSO interactions. Similarly, reference [11] proposes a model to operate transmission and distribution systems in a coordinated manner using a surrogate Lagrangian Relaxation approach. Finally, an analytical target cascading procedure (ATC) to coordinate the operation of transmission and distribution networks is described in [12].

The decomposition and decentralized methods previously described are able to obtain the same solution as the centralized approach while significantly reducing the computational burden. Yet these methods also have meaningful drawbacks.

Decomposition methods still require full access to all physical and economic information on distribution networks. However, as stated in [13] "distribution system operators are autonomous entities that are unwilling to reveal their commercially sensitive information." Therefore, these methods can hardly be accommodated in a real-life distribution environment with even a few ambiguous or unknown parameters (e.g. topological configuration, impedance, voltage and flow limits) and proprietary customer-end and behind-the-meter parameters (e.g. production/utility cost functions, supply/demand elasticity and behavioral aspects of electricity demand). Similarly, decentralized methods are based on repetitive real-time information exchanges between the TSO and the DSOs, and thus, rely on robust and fast communication infrastructure. As discussed in [9], "The communication infrastructures for implementing distributed methods need to be carefully designed, and the impact of communication delays and failures on the performance of distributed methods need to be investigated".

Instead of using decomposition or decentralized procedures, which heavily rely on access to either all physical and economic information or fail-safe real-time communication infrastructure, what we propose in this paper is an approximate method that requires neither of these two controversial assumptions at the expense of obtaining a solution slightly different from the optimal one. The proposed approach only needs access to offline historical information on prices and power injection at the substations connecting the transmission and distribution networks. Using statistical tools, we learn the price response of active distribution networks whose operating decisions aim at minimizing costs while complying with local physical constraints, such as voltage limits. We also utilize easily accessible information (e.g., capacity factors of wind and solar local resources) to make the curve adaptive to changes in external conditions that affect power system operations. Finally, the obtained response is approximated by a non-increasing step-wise function that can be conveniently interpreted as a market bid for the participation of the distribution networks in wholesale electricity markets. In summary, the contributions of our paper are twofold:

1) We propose a TSO-DSO coordination scheme that uses historical data at substations to learn the price response of active distribution networks using a decreasing step-wise function. If compared with existing methodologies, ours is simple and easy to implement within current market procedures, and cheap in terms of computational resources, information exchange and communication infrastructure.

2) We measure the performance of the proposed approach in terms of the power imbalances and social welfare loss caused by the approximation of the distribution networks' behavior using a realistic case study.

We compare our approach against a fully centralized operational model, referred to as *benchmark*, that guarantees the optimal coordination between the TSO and the DSOs. Since this benchmark produces the same solution obtained with exact decomposition and decentralized methods [9]–[13], these have not been considered in our study. On the contrary, our model
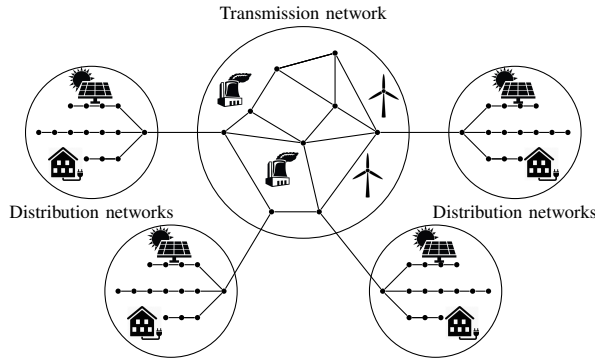
Fig. 1. Transmission and distribution coordination scheme

Fig. 2. Flexible electricity demand modeling

is evaluated against two other approximations. In the first one, called the *single-bus approach*, all physical constraints of the distribution networks are disregarded as if all small consumers and distributed generating resources were directly connected to the main substation. In the second one, called the *price-agnostic approach*, the response of the distribution networks is assumed to be independent of prices.

We note that, within the context of reactive power optimization for the minimization of network losses, the authors in [14] also approximate the apparent power exchange between the TSO and the DSOs by a polynomial function of the voltage level at the main substation. However, beyond the evident facts that their purpose is different and the fitting procedure we need to use is more intricate (to comply with market rules), they also omit the *dynamic* nature of distribution network response to local marginal prices (LMPs). Similarly, authors of [15] also propose a methodology to obtain the relation between the distribution network response and the voltage level at the substation. However, their approach is based on perfect knowledge of all distribution network parameters.

The rest of this paper is organized as follows. Section II introduces optimization models for transmission and distribution network operations, which are then used to construct different DSO-TSO coordination approaches in Section III. The metrics we use for comparing these approaches are described in Section IV, while the case study is presented in Section V. Finally, conclusions are duly reported in Section VI.

II. MODELING FRAMEWORK

We consider a power system with a high-voltage, meshed transmission network connected to generating units, large consumers and several medium-voltage distribution networks. As illustrated in Fig. 1, each distribution system is connected to the transmission network through one main substation, has a radial topology and hosts small-scale electricity consumers and producers.

The active power output of generating unit i at time period t is denoted by p_{it}^G, with minimum/maximum limits $\underline{p}_i^G/\overline{p}_i^G$. Generating units are assumed to have a convex cost function of the form $c_i(p_{it}^G) = \frac{1}{2}a_i(p_{it}^G)^2 + b_i(p_{it}^G)$, with $a_i, b_i \geq 0$, and a dimensionless capacity factor ρ_{it}, with $0 \leq \rho_{it} \leq 1$. For thermal units $\rho_{it} = 1, \forall t$, while for renewable generating units the capacity factor depends on weather conditions and the production cost is zero ($a_i = b_i = 0$).
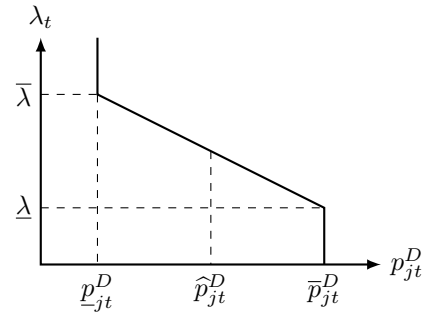
Electricity consumption is modeled as a capped linear function of the LMP λ_t, as shown in Fig. 2, where $\widehat{p}_{jt}^D$ denotes the baseline demand of consumer j at time t and $\overline{p}_{jt}^D/\underline{p}_{jt}^D$ are the maximum/minimum load levels given by $\overline{p}_{jt}^D = \widehat{p}_{jt}^D(1+\delta_j)$ and $\underline{p}_{jt}^D = \widehat{p}_{jt}^D(1-\delta_j)$, with $\delta_j \geq 0$ [16]. Consequently, the maximum/minimum load levels vary over time according to the evolution of the baseline demand. Under this modeling approach, a price-insensitive demand is modeled with $\delta_j = 0$, while $\delta_j = 0.5$ implies that the consumer is willing to increase or decrease their baseline demand up to 50% depending on the price. Finally, $\overline{\lambda}$ and $\underline{\lambda}$ stand for the LMP values that unlock the minimum and maximum demand from consumers, respectively. The demand function in Fig. 2 goes through points $(\underline{p}_{jt}^D, \overline{\lambda})$ and $(\overline{p}_{jt}^D, \underline{\lambda})$ and, therefore, its expression can be determined as follows:

$$\frac{\lambda_t - \underline{\lambda}}{\overline{\lambda} - \underline{\lambda}} = \frac{p_{jt}^D - \overline{p}_{jt}^D}{\underline{p}_{jt}^D - \overline{p}_{jt}^D} = \frac{p_{jt}^D - \widehat{p}_{jt}^D(1+\delta_j)}{\widehat{p}_{jt}^D(1-\delta_j) - \widehat{p}_{jt}^D(1+\delta_j)} \implies$$

$$p_{jt}^D = \widehat{p}_{jt}^D\left(1 + \delta_j\frac{\overline{\lambda}+\underline{\lambda}}{\overline{\lambda}-\underline{\lambda}}\right) - \frac{2\widehat{p}_{jt}^D\delta_j}{\overline{\lambda}-\underline{\lambda}}\lambda_t \qquad (1)$$

Hence, the active demand level p_{jt}^D for a given electricity price λ_t takes the following form:

$$p_{jt}^D = \begin{cases} \overline{p}_{jt}^D & \text{if} \quad \lambda_t \leq \underline{\lambda} \\ \alpha_{jt} - \beta_{jt}\lambda_t & \text{if} \quad \underline{\lambda} < \lambda_t < \overline{\lambda} \\ \underline{p}_{jt}^D & \text{if} \quad \overline{\lambda} \leq \lambda_t, \end{cases} \qquad (2)$$

where $\alpha_{jt} = \widehat{p}_{jt}^D\left(1 + \delta_j\frac{\overline{\lambda}+\underline{\lambda}}{\overline{\lambda}-\underline{\lambda}}\right)$ and $\beta_{jt} = \frac{2\widehat{p}_{jt}^D\delta_j}{\overline{\lambda}-\underline{\lambda}}$. The reactive power demand is given by $q_{jt}^D = \gamma_j p_{jt}^D$, where γ_j is the power factor of consumer j, which is assumed to be independent of time for simplicity. Finally, we obtain the utility of each consumer by integrating the inverse demand function with respect to the demand quantity, that is,

$$u_{jt}(p_{jt}^D) = \int_{\underline{p}_{jt}^D}^{p_{jt}^D} \lambda(p)dp = \int_{\underline{p}_{jt}^D}^{p_{jt}^D} \left(\frac{\alpha_{jt}}{\beta_{jt}} - \frac{p}{\beta_{jt}}\right)dp =$$

$$= \frac{\alpha_{jt}}{\beta_{jt}}\left(p_{jt}^D - \underline{p}_{jt}^D\right) - \frac{(p_{jt}^D)^2 - (\underline{p}_{jt}^D)^2}{2\beta_{jt}} \qquad (3)$$

The transmission network is modeled using a DC power flow approximation [17] and, therefore, each line l going from node o_l to node e_l is characterized by its reactance x_l and maximum capacity $\overline{s}_l^F$. The power flow is denoted by p_{lt}^F.

Then, suppose that the active consumption of the k-th distribution network p_{kt}^N can be expressed as $p_{kt}^N = h_{kt}(\lambda_{kt})$, where λ_{kt} is the price at the corresponding substation. Under this assumption, transmission system operations at time period t are modeled by the following optimization problem:

$$\max_{\Phi_t^T} \sum_{j \in J^T} u_{jt}(p_{jt}^D) + \sum_{k \in K^T} \int_0^{p_{kt}^N} h_{kt}^{-1}(s)ds - \sum_{i \in I^T} c_i(p_{it}^G) \quad (4a)$$

s.t.

$$\sum_{i \in I_n} p_{it}^G - \sum_{j \in J_n} p_{jt}^D - \sum_{k \in K_n} p_{kt}^N =$$
$$= \sum_{l:e_l=n} p_{lt}^F - \sum_{l:o_l=n} p_{lt}^F, \ \forall n \in N^T \quad (4b)$$

$$\frac{p_{lt}^F}{s^B} = \frac{1}{x_l}(\theta_{o_l t} - \theta_{e_l t}), \ \forall l \in L^T \quad (4c)$$

$$\underline{p}_i^G \leq p_{it}^G \leq \rho_{it}\overline{p}_i^G, \ \forall i \in I^T \quad (4d)$$

$$\underline{p}_{jt}^D \leq p_{jt}^D \leq \overline{p}_{jt}^D, \ \forall j \in J^T \quad (4e)$$

$$-\overline{s}_l^F \leq p_{lt}^F \leq \overline{s}_l^F, \ \forall l \in L^T \quad (4f)$$

where θ_{nt} is the voltage angle at node n and time period t, $\Phi_t^T = (p_{it}^G, p_{jt}^D, p_{kt}^N, p_{lt}^F, \theta_{nt})$ are decisions variables, N^T, L^T, I^T, J^T, K^T are sets of nodes, lines, generators, consumers and distribution networks connected to the transmission network, and I_n, J_n, K_n are sets of generating units, consumers and distribution networks connected to node n. Objective function (4a) maximizes the total social welfare and includes the utility of all flexible consumers connected to the transmission network (first term), the utility of all distribution networks (second term), and the generation cost of all units connected to the transmission network (third term). Note that $h_{kt}^{-1}(\cdot)$ represents the inverse demand function and its integral correspond to the total utility of each distribution network. The nodal power balance equation is imposed by (4b), while the power flow through each transmission line is computed in (4c). Finally, constraints (4d), (4e) and (4f) enforce the generation, consumption and transmission capacity limits.

Traditionally, distribution networks only hosted inflexible consumption and, therefore, p_{kt}^N was considered independent of the electricity price. In this case, the second term of (4a) vanishes, and variable p_{kt}^N is replaced by the forecast power intake of each distribution network. Thus, problem (4) can be transformed into a quadratic optimization problem that can be solved to global optimality using off-the-shelf solvers, [18, Appendix B]. However, this paradigm has changed in recent years and current distribution networks include a growing amount of flexible small-scale consumers and distributed generation resources that are capable of adjusting their consumption/generation in response to the electricity price to maximize their utility/payoff [19]. Indeed, if λ_{kt} is the electricity price at the main substation of distribution network k, the power intake of that distribution network p_{kt}^N can be determined by solving the following optimization problem:

$$\max_{\Phi_{kt}^D} \sum_{j \in J_k^D} u_{jt}(p_{jt}^D) - \sum_{i \in I_k^D} c_i(p_{it}^G) - \lambda_{kt}p_{kt}^N \quad (5a)$$

s.t.

$$p_{kt}^N + \sum_{i \in I_n} p_{it}^G - \sum_{j \in J_n} p_{jt}^D =$$
$$= \sum_{l:e_l=n} p_{lt}^F - \sum_{l:o_l=n} p_{lt}^F, \ n = n_k^0 \quad (5b)$$

$$\sum_{i \in I_n} p_{it}^G - \sum_{j \in J_n} p_{jt}^D =$$
$$= \sum_{l:e_l=n} p_{lt}^F - \sum_{l:o_l=n} p_{lt}^F, \ \forall n \in N_k^D, n \neq n_k^0 \quad (5c)$$

$$\sum_{i \in I_n} q_{it}^G - \sum_{j \in J_n} q_{jt}^D =$$
$$= \sum_{l:e_l=n} q_{lt}^F - \sum_{l:o_l=n} q_{lt}^F, \ \forall n \in N_k^D \quad (5d)$$

$$q_{jt}^D = \gamma_j p_{jt}^D, \ \forall j \in J_k^D \quad (5e)$$

$$v_{nt} = v_{a_n t} - \frac{2}{s^B} \sum_{l:e_l=n} r_l p_{lt}^F + x_l q_{lt}^F, \ \forall n \in N_k^D \quad (5f)$$

$$\underline{p}_i^G \leq p_{it}^G \leq \rho_{it}\overline{p}_i^G, \ \forall i \in I_k^D \quad (5g)$$

$$\underline{q}_i^G \leq q_{it}^G \leq \overline{q}_i^G, \ \forall i \in I_k^D \quad (5h)$$

$$(p_{it}^G)^2 + (q_{it}^G)^2 \leq (\overline{s}_i^G)^2, \ \forall i \in I_k^D \quad (5i)$$

$$\underline{p}_{jt}^D \leq p_{jt}^D \leq \overline{p}_{jt}^D, \ \forall j \in J_k^D \quad (5j)$$

$$(p_{lt}^F)^2 + (q_{lt}^F)^2 \leq (\overline{s}_l^F)^2, \ \forall l \in L_k^D \quad (5k)$$

$$\underline{v}_{nt} \leq v_{nt} \leq \overline{v}_{nt}, \ \forall n \in N_k^D \quad (5l)$$

where the decisions variables are $\Phi_{kt}^D = (p_{kt}^N, p_{it}^G, q_{it}^G, p_{jt}^D, q_{jt}^D, p_{lt}^F, q_{lt}^F, v_{nt})$. In particular, $q_{it}^G, q_{jt}^D, q_{lt}^F$ are the reactive power generation, consumption and flow, in that order, and v_{nt} is the squared voltage magnitude. Since we assume a radial distribution network, we use the LinDistFlow AC power flow approximation, where a_n represents the ancestor of node n and r_l is the resistance of line l [20]. The rate power of the inverters for distributed generators is denoted as $\overline{s}_i^G$ [21], and the squared voltage magnitude limits are $\underline{v}_{nt}, \overline{v}_{nt}$. Finally, $N_k^D, L_k^D, I_k^D, J_k^D$ are the set of nodes, lines, generators and consumers of distribution network k, and n_k^0 corresponds to the node of the distribution network connected to the substation.

Objective function (5a) maximizes the social welfare of distribution network k and includes the utility of flexible consumers (first term), the cost of distributed generation (second term) and the cost of power exchanges with the transmission network (third term). Nodal active and reactive power equations are formulated in (5b), (5c) and (5d). Constraint (5e) relates active and reactive demand through a given power factor, while the dependence of voltage magnitudes in a radial network is accounted for in (5f) using the LinDistFlow approximation. Limits on active and reactive generating power outputs are enforced in (5g), (5h) and (5i). Similarly, equations (5j), (5k) and (5l) determine the feasible values of demand quantities, power flows and squared voltage magnitudes. As a result, (5) is a convex optimization problem that can be solved using off-the-shelf solvers.

Drawing a closed-form expression $h_{kt}(\lambda_{kt})$ from (5) that exactly characterizes the optimal value of p_{kt}^N as a function of the electricity price λ_{kt} seems like a lost cause. Furthermore, even if such an expression were possible, using it in (4) would

lead to a troublesome non-convex optimization problem, with the likely loss of global optimality guarantees. In the next section, we discuss different strategies to construct an approximation $\hat{h}_{kt}(\lambda_{kt})$ that can be easily incorporated into (4) to determine the optimal operation of the transmission network. In particular, we focus on strategies that leverage available contextual information to construct function $\hat{h}_{kt}(\lambda_{kt})$.

III. METHODOLOGY

In this section we present four different approaches to accommodate the behavior of active distribution networks in transmission network operations. The aim of these methods is to determine the electricity prices at substations that foster the most efficient use of the flexible resources available in the distribution networks. Once electricity prices are published, the DSOs operate the distributed energy resources to maximize their social welfare while satisfying the physical limits of the distribution network, such as voltage limits and reactive power capacities.

A. Benchmark approach (BN)

This approach includes a full representation of both the transmission system and the distribution networks, by jointly solving optimization problems (4) and (5) as follows:

$$\max_{\Phi_t^T, \Phi_{kt}^D} \sum_{j \in J^T \cup \{J_k^D\}} u_{jt}(p_{jt}^D) - \sum_{i \in I^T \cup \{I_k^D\}} c_i(p_{it}^G) \tag{6a}$$

$$\text{s.t.} \quad (4b) - (4f), (5c) - (5l) \tag{6b}$$

Model (6) enables the optimal operation of the transmission network since it takes into account the most accurate representation of all distribution networks connected to it [22]. However, this approach has the following drawbacks:

- It requires having access to distribution network parameters, such as its topological configuration and r_l, x_l, which is impractical, as private or sovereign entities operating distribution networks prefer to keep this information confidential [13], [14], [23].
- Operating the power system through (6) would require a deep transformation of current market mechanisms to allow small generators/consumers to directly submit their electricity offers/bids to a centralized market operator.
- Even if all distribution network parameters were known and small generators/consumers were allowed to directly participate in the electricity market, solving model (6) is computationally expensive for realistically sized systems with hundreds of distribution networks connected to the transmission network [2].

In this paper, we use the solution of this approach as a benchmark to evaluate the performance of the other methods described in this section. Other decomposition or decentralized methods in the technical literature are able to achieve global optimality and, therefore, their solution coincide with that of BN. For this reason, we focus on comparing the proposed approach with other approximate methodologies that also lead to suboptimal solutions.

B. Single-bus approach (SB)

This approach is a relaxation of BN in (6), where physical limits on distribution power flows and voltages are disregarded. Therefore, operational model SB can be equivalently interpreted as if all small consumers and distributed energy resources were directly connected to the transmission network, i.e. all distribution systems are modeled as single-bus grids. Therefore, the dispatch decisions for the transmission network are computed by solving the following problem:

$$\max_{\Phi_t^T, \Phi_{kt}^D} \sum_{j \in J^T \cup \{J_k^D\}} u_{jt}(p_{jt}^D) - \sum_{i \in I^T \cup \{I_k^D\}} c_i(p_{it}^G) \tag{7a}$$

s.t.

$$\sum_{i \in \hat{G}_n} p_{it}^G - \sum_{j \in \hat{D}_n} p_{jt}^D = \sum_{l:e_l=n} p_{lt}^F - \sum_{l:o_l=n} p_{lt}^F, \ \forall n \in N^T \tag{7b}$$

$$\frac{p_{lt}^F}{s^B} = \frac{1}{x_l}(\theta_{o_l t} - \theta_{e_l t}), \ \forall l \in L^T \tag{7c}$$

$$\underline{p}_i^G \le p_{it}^G \le \rho_{it}\overline{p}_i^G, \ \forall i \in I^T \cup \{I_k^D\} \tag{7d}$$

$$\underline{p}_{jt}^D \le p_{jt}^D \le \overline{p}_{jt}^D, \ \forall j \in J^T \cup \{J_k^D\} \tag{7e}$$

$$-\overline{s}_l^F \le p_{lt}^F \le \overline{s}_l^F, \ \forall l \in L^T \tag{7f}$$

where $\hat{G}_n$ and $\hat{D}_n$ denote, respectively, the set of generators and consumers either directly connected to node n or hosted by a distribution network connected to it. Problem (7) is less computationally demanding than the BN approach in (6) and does not require knowledge of distribution network parameters. However, this approach also relies on a market mechanism that allows small generators and consumers to submit their offers and bids directly to the wholesale market [24]. Besides, if the operation of some of the distribution networks is constrained by the physical limitations of power flows and/or voltage levels, then the solution provided by this approach may substantially differ from the actual conditions in the distribution networks.

C. Contextual price-agnostic approach (PAG)

This approach is based on the premise that the penetration rates of small-scale flexible consumers and distributed generation resources is not significant and, therefore, the response of distribution networks is independent of LMPs at their substations. On the other hand, this response can still depend on other contextual information that affect the behavior of distribution networks such as the aggregated load level of their flexible consumers and the wind and solar capacity factors in the corresponding geographical area.

Consider a set of historical data $\{\chi_{kt}, p_{kt}^N\}_{t \in \mathcal{T}}$, where χ_{kt} represents a vector containing the contextual information to explain the consumption level of distribution network k. Vector χ_{kt} can include weather conditions, e.g. ambient temperature, wind speed, solar irradiation, or categorical variables, e.g. an hour of the day or a day of the week. The PAG approach aims to learn the relation between p_{kt}^N and χ_{kt} for each distribution network k, i.e.,

$$p_{kt}^N = f_k(\chi_{kt}) \tag{8}$$

The function f_k that best approximates the behavior of distribution network k with contextual information can be found

using a wide variety of supervised learning techniques [25]. In particular, if f_k must belong to a certain family of functions, such as the family of linear functions, its parameters can be computed using the well-known least squares criterion. In order to capture non-linear relations, f_k can also represent a neural network to be trained using available data. Alternatively, if we do not make strong assumptions about the form of the mapping function, the relation between χ_{kt} and p_{kt}^N can be modeled using non-parametric supervised learning techniques. Within this group, we opt in this work for the k-nearest neighbors regression algorithm (K-NN) because of its simplicity, interpretability and scalability. Following this methodology, the estimation of the power import of distribution network k for time period t (denoted as $\widehat{p}_{kt}^N$) is computed as:

$$\widehat{p}_{kt}^N = \frac{1}{K} \sum_{t' \in \mathcal{T}^C(t)} p_{kt'}^N, \qquad (9)$$

where $\mathcal{T}^C(t)$ is the subset of the K time periods whose contexts are the closest to χ_{kt} according to a given distance, and t' is an auxiliary time period index. If contextual information only includes continuous variables (electricity demand, renewable power generation, etc.), the dissimilarity between two time periods can be measured using the Euclidean distance, i.e., $dist(t_1, t_2) = ||\chi_{kt_1} - \chi_{kt_2}||_2$. If contextual information also includes binary variables (equipment status, maintenance schedules, etc.), the dissimilarity can be measured using the Hamming distance, for example.

Once the forecast intake for each distribution network is obtained depending on its corresponding context, we model all distribution networks as fix loads and determine the operation of the transmission network by solving the following optimization problem:

$$\max_{\Phi^T} \sum_{j \in J^T} u_{jt}(p_{jt}^D) - \sum_{i \in I^T} c_i(p_{it}^G) \qquad (10a)$$

s.t.

$$p_{kt}^N = \widehat{p}_{kt}^N, \ \forall k \in K \qquad (10b)$$

$$(4b) - (4f) \qquad (10c)$$

Problem (10) is also more computationally tractable than (6) and does not rely on knowledge of distribution network parameters. As the one we propose, this approach only requires access to historical power flow measurements at the substation and contextual information that can enhance explainability and interpretability of the distribution network responses to external factors of interest. Fortunately, independent system operators such as ISONE and NYISO make this information publicly available. Another advantage of this approach is that, unlike the SB approach, it can be seamlessly implemented in existing market-clearing procedures since the response of distribution networks is simply replaced with the fixed power injections provided by (9). Actually, this is the approach that most closely reproduces the traditional way of proceeding. On the other hand, since the impact of substation LMPs on the response of distribution networks is disregarded, the accuracy of this approach worsens as the flexibility provided by small consumers and distributed generators increase.
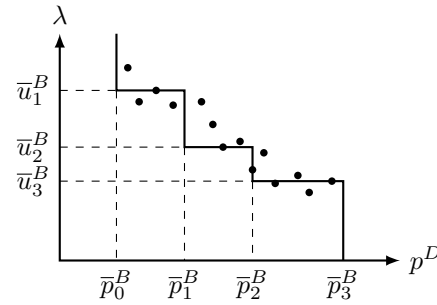

Fig. 3. Step-wise approximation of distribution network response

D. Contextual price-aware approach (PAW)

The SB and PAG approaches disregard the impact of either physical limits or economic signals on the response of distribution networks with small-scale, flexible consumers and distributed generation resources. To overcome this drawback, we propose to approximate the response function $h_{kt}(\lambda_{kt})$ by taking into account the effects of both physical and economic conditions on the behavior of active distribution networks.

Similarly to the PAG approach, we assume access to the set of historical data $\{\chi_{kt}, \lambda_{kt}, p_{kt}^N\}_{t \in \mathcal{T}}$, where λ_{kt} denotes the LMP at the substation of distribution network k. The proposed PAW approach aims at determining the function that explains the response p_{kt}^N as a function of the contextual information χ_{kt} and the electricity price at the substation λ_{kt}, i.e.,

$$p_{kt}^N = g_k(\chi_{kt}, \lambda_{kt}) \qquad (11)$$

For a fixed context, function (11) provides the relation between the response of a distribution network and the price at its substation. This function can be understood as the bid to be submitted by each distribution network to the wholesale electricity market. However, most current market procedures only accept a finite number of decreasing block bids. For instance, the Spanish market operator establishes that "For each hourly scheduling period within the same day-ahead scheduling horizon, there can be as many as 25 power blocks for the same production unit, with a different price for each of the said blocks, with the prices increasing for sale bids, or decreasing for purchase bids." [26]. In order to comply with these market rules, we propose an efficient learning procedure to determine a decreasing step-wise mapping between the response of a distribution network and the LMP, while taking into account contextual information. Our approach combines unsupervised and supervised learning techniques as follows:

- *Unsupervised learning*. Similarly to PAG, the first step of the proposed approach uses a K-nearest neighbors algorithm to find the subset of time periods $\mathcal{T}^C(t)$ whose contextual information are the closest to χ_{kt}. For the sake of illustration, the points depicted in Fig. 3 represent the pairs of prices and power intakes for times periods in $\mathcal{T}^C(t)$ for a given substation k and context χ_{kt}.
- *Supervised learning*. The second step of the proposed approach consists of finding the step-wise decreasing function that best approximates the price-quantity pairs obtained in the previous step. As illustrated in Fig. 3,
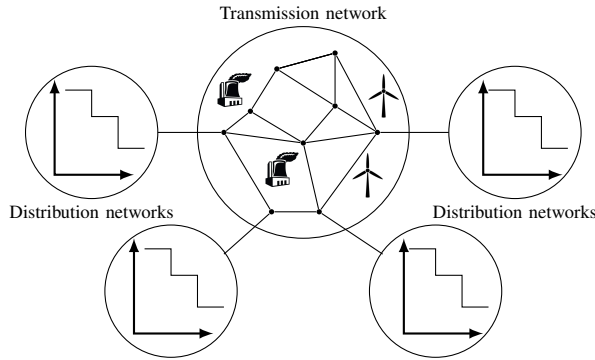
Fig. 4. Proposed TSO-DSO coordination approach

this function can be defined by a set of price break-points $\overline{u}_b^B$ and the demand level for each block $\overline{p}_b^B$. Despite its apparent simplicity, finding the optimal step-wise decreasing function that approximates a set of data points is a complex task that cannot be accomplished by conventional regression techniques. For instance, isotonic regression yields a monotone step-wise function, but a maximum number of blocks cannot be imposed. Conversely, segmented regression provides a step-wise function with a maximum number of blocks, but monotonicity is not ensured. Therefore, the statistical estimation of $\overline{p}_0^B$ and $\overline{u}_b^B$ and $\overline{p}_b^B$, $\forall b \in \mathcal{B}$, is conducted by means of the curve-fitting algorithm for segmented isotonic regression that has been recently developed in [27] and can be formulated as the following optimization problem:

$$\min_{\overline{p}_0^B, \overline{u}_b^B, \overline{p}_b^B} \sum_{t' \in \mathcal{T}^C(t)} \left(p_{kt'}^N - \sum_{b \in \mathcal{B} \cup \{0\}} \overline{p}_b^B \mathbb{I}_{[\overline{u}_{b+1}^B, \overline{u}_b^B)}(\lambda_{kt'}) \right)^2 \quad (12a)$$

$$\text{s.t.} \quad \overline{p}_b^B \geq \overline{p}_{b-1}^B, \quad \forall b \in B \quad (12b)$$

$$\overline{u}_{b+1}^B \leq \overline{u}_b^B, \quad \forall b \in B \quad (12c)$$

where $\mathbb{I}_{[\overline{u}_{b+1}^B, \overline{u}_b^B)}$ is the indicator function equal to 1 if $\overline{u}_{b+1}^B \leq \lambda_{kt} < \overline{u}_b^B$, and 0 otherwise, and $\overline{u}_0^B = \infty$, $\overline{u}_{|\mathcal{B}|+1}^B = -\infty$. Objective function (12a) minimizes the sum of squared errors, while constraints (12b)-(12c) ensures the monotonicity of the regression function. Problem (12) can be reformulated as a mixed-integer quadratic problem to be solved by standard optimization solvers. However, the computational burden of this solution strategy is extremely high. Alternatively, reference [27] proposes a dynamic programming reformulation that guarantees global optimality in polynomial time, which makes this approach computationally attractive.

In summary, we approximate the response of the distribution networks using a learning strategy that combines a K-nearest neighbor algorithm and a curve-fitting methodology. The proposed learning approach is simple and fast, as well as it does not suffer from high data requirements for training and offers explainability of the results, unlike black-box approaches, e.g. based on deep learning. The operation of the transmission network is obtained by assuming that each distribution network reacts to prices according to the obtained step-wise non-increasing functions as illustrated in Fig. 4. Mathematically,

TABLE I
Qualitative comparison of TSO-DSOs coordination approaches

	BN	SB	PAG	PAW
Network-aware	X		X	X
Price-aware	X	X		X
Historical data			X	X
Seamless market integration			X	X
Computational burden	High	Low	Low	Low

the operation of the transmission network is obtained by solving the following optimization problem:

$$\max_{\Phi_t^T, p_{kt}^N, p_{ktb}^B} \sum_{b \in \mathcal{B}, k \in K^T} \overline{u}_{ktb}^B p_{ktb}^B + \sum_{j \in J^T} u_{jt}(p_{jt}^D) - \sum_{i \in I^T} c_i(p_{it}^G)$$
$$(13a)$$

$$\text{s.t.} \quad p_{kt}^N = \overline{p}_{kt0}^B + \sum_{b \in \mathcal{B}} p_{ktb}^B, \ \forall k \in K^T \quad (13b)$$

$$0 \leq p_{ktb}^B \leq \overline{p}_{ktb}^B - \overline{p}_{kt(b-1)}^B, \ \forall b \in \mathcal{B}, k \in K^T \quad (13c)$$

$$(4b) - (4f) \quad (13d)$$

The proposed approach has several advantages. First, while the SB and PAG approaches disregard, respectively, the impact of network limits or economic signals on the response of distribution networks, the PAW approach is aware of both effects. Second, like the PAG approach, this method only requires historical LMPs and power flows at the substations and, therefore, detailed information about the distribution network parameters is not required. Third, the response of each distribution network to prices is modeled by a step-wise decreasing function that can be directly included in existing market-clearing mechanisms without additional modifications. Besides, unlike other decomposition/decentralized approaches, the one we propose is not an iterative method and, therefore, is immune to convergence issues. Since the proposed method basically relies on a learning task, its performance highly depends on the quality of the input historical data. Hence, the dataset should be continuously updated to include the most recent operating conditions and exclude the oldest ones.

To conclude this section, Table I summarizes the main features of the four approaches discussed above. If compared with the benchmark, the three alternative approaches involve lower computational burdens through different approximation strategies. The next section describes the methodology to quantify the impact of such approximations on the optimal operation of the transmission electricity network.

IV. EVALUATION PROCEDURE

While existing decomposition/decentralized approaches are able to yield the optimal coordination decisions, the proposed methodology may lead to suboptimal decisions caused by incompleteness or inaccuracies of the learning task. In this section, we present the evaluation procedure to quantify the impact of these suboptimal decisions in terms of power imbalance and social welfare losses. We compare such measures with those obtained by the other three methods described in Section III. To that end, we proceed as follows:

1) Solve problems (6), (7), (10) or (13) using the modeling of the distribution networks derived from the BN, SB, PAG or PAW approaches. LMPs at each substation λ_{kt} are obtained as the dual variable of the balance equation (4b). The sum of the approximated consumption by all distribution networks is denoted as $\widehat{P}_t^N$.

2) Model (5) is solved for each distribution network k after fixing LMPs at the substations to those obtained in Step 1). As such, we compute the actual response of the distribution networks considering all physical and economic information, denoted as P_t^N. Optimal values of objective function (5a) provide the social welfare achieved by each distribution network for the electricity prices computed in Step 1). We denote the sum of the social welfare of all distribution networks as SW_t^D.

3) Quantify the power imbalance caused by the different distribution network approximations as $\Delta_t = 100|\widehat{P}_t^N - P_t^N|/P_t^N$. Note that such power imbalances must be handled by flexible power resources able to adapt their generation or consumption in real-time.

4) Model (4) is solved by setting the electricity imported by each distribution network to the quantity obtained in Step 2). The output of this model represents the real-time re-dispatch of generating units connected to the transmission network to ensure the power system balance. The optimal value of (4a) provides the realized social welfare of the transmission network denoted as SW_t^T. We emphasize that this social welfare is computed as if all generating units and consumers at the transmission network could instantly adapt to any unexpected power imbalance coming from the distribution networks (Δ_t) without any extra cost for the deployment of such unrealistic flexible resources. This means that we are underestimating the social welfare loss caused by these power imbalances.

5) Compute the total realized social welfare of the power system as $SW_t = SW_t^D + SW_t^T$.

V. SIMULATION RESULTS

We consider the 118-bus, 186-line transmission network from [28]. Each transmission-level load is replaced with a 32-bus radial distribution network, which hosts eight solar generating units, see data in [29], [30]. That is, the power system includes 3030 buses ($118 + 91 \times 32$), 3098 lines ($186 + 91 \times 32$), thermal and wind power plants connected to 43 transmission buses, solar generating units connected to 728 distribution buses (91×8), and electricity consumers located at 2912 distribution buses (91×32). Each consumer is assumed to react to the electricity price as depicted in Fig. 2. The installed capacity of thermal, solar and wind generating units is 17.3GW, 2.5GW and 2.5GW, respectively, while the peak demand is 18GW. Finally, time-varying capacity factors for all consumers, wind and solar generation in the same distribution network are assumed equal. While all distribution networks have the same topology and the same location of loads and solar power generating units, we scale their total demand from 12MW to 823MW to match the transmission

demand given in [28]. We also scale the original values of branch resistances and reactances inversely proportional to the peak demand within each distribution network. All data used in this case study is available in [31]. Simulations have been run on a Linux-based server with one CPU clocking at 2.6 GHz and 2 GB of RAM using CPLEX 12.6 under Pyomo 5.2.

As discussed in Section III, the analyzed methods differ in their ability to account for the impact of physical limits and economic signals on the response of active distribution networks. For instance, if distribution voltage limits never become activated, then the SB approach would provide results quite close to those of the benchmark approach BN. Conversely, if distribution voltages reach their security limits, the PAG and PAW methods are expected to outperform SB. In order to investigate the impact of voltage congestion on the performance of each approach, we vary the resistances and reactances of branches of the distribution networks as indicated in (14), where r^0, x^0 are the base-case values provided in [31], and parameter η is changed from 0.67 to 1.33, i.e., a 33% lower and greater than the initial values:

$$r = \eta r^0 \qquad x = \eta x^0 \qquad (14)$$

Additionally, we use parameter δ to model each flexible consumer, which is randomly generated for the 2912 loads following a uniform probability distribution in $[0.5 - 0.75]$. Besides, we set $\overline{\lambda} = 25$, $\underline{\lambda} = 10$, $\overline{v}_n = 1.05$, and $\underline{v}_n = 0.95$.

The PAG and PAW approaches require access to historical data. In this case study, historical data is generated by solving the BN model (6) for 8760 hours of a given year. Each hour is characterized by different baseline demands of flexible consumers along with the wind and solar capacity factors throughout the system. Values for wind, solar and baseline demands are taken from [28] and are available in [31]. For illustration, Fig. 5 plots the price response of one distribution network for $\eta = 1$ including the 8760 time periods. For simplicity, changes in the topology of the transmission and distribution networks are disregarded. The learning-based PAG and PAW approaches use the demand and renewable capacity factors at each distribution network as contextual information to learn its response. Also, the number of neighbors for the K-NN learning methodology is set to 100. Finally, the maximum number of blocks for the bidding curves learned by the PAW approach is equal to ten. For the sake of comparison, each of the four approaches uses the same test set that includes 100 randomly selected hours of the year.

Using the results of these 100 hours, Fig. 6 plots, for each approach, a shaded area ranging from the 5% to the 95% percentile of the relative power imbalance Δ_t as a function of parameter η. The average of the power imbalance is also displayed. Naturally, due to its completeness, the benchmark method does not incur any power imbalance and therefore, the results delivered by this method are not included in Fig. 6. Low values of η reduce voltage congestion at the distribution networks and, therefore, their response is mainly driven by electricity prices at the substations. In such cases, the SB approach outperforms the PAG approach and yields power imbalances close to 0%. For small values of η, the proposed PAW approach yields higher power imbalances than
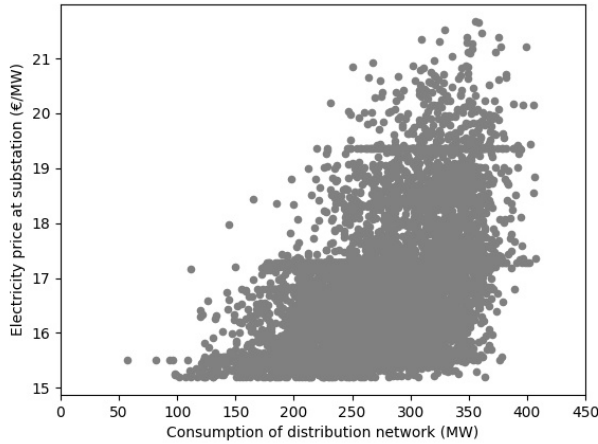
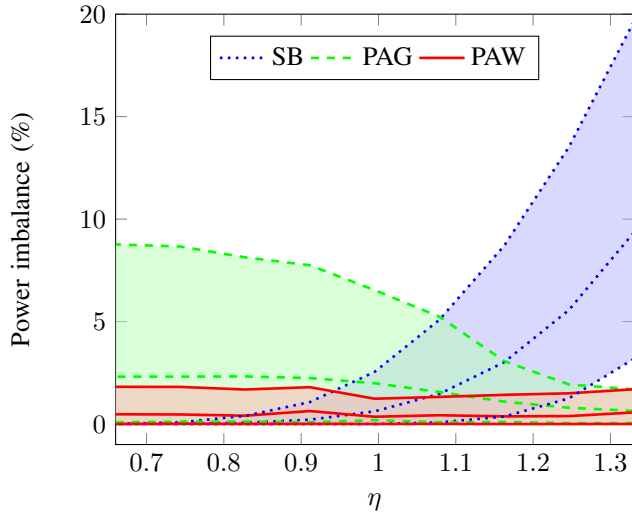Fig. 5. Price response of one distribution network for $\eta = 1$

Fig. 6. Impact of distribution network congestion on power imbalance.

Fig. 7. Estimated (dashed) and observed (bold) flexibility of active distribution networks.

SB. However, this difference could be narrowed by approximating the response of the distribution networks with more than ten blocks. Conversely, high values of η translates into congested distribution networks in which the dispatch of small consumers and distributed generators is heavily constrained by technical limits. In these circumstances, electricity prices at the substations have a reduced impact on the the response of the distribution network and then, the power imbalance of the SB approach is significantly greater than that of the PAG approach. Quantitatively, the proposed methodology PAW achieves average power imbalances below 0.7% for any value of η.

When comparing SB, PAG and PAW, we shoud also keep in mind that their integration into current market-clearing mechanisms are not comparable. Implementing the SB approach would require modifying existing market rules so that distributed generators and small consumers could directly submit their offers and bids. On the other hand, the PAG and PAW comply with these rules since active distribution networks are modeled as fix loads or in the form of step-wise bidding curves, respectively.
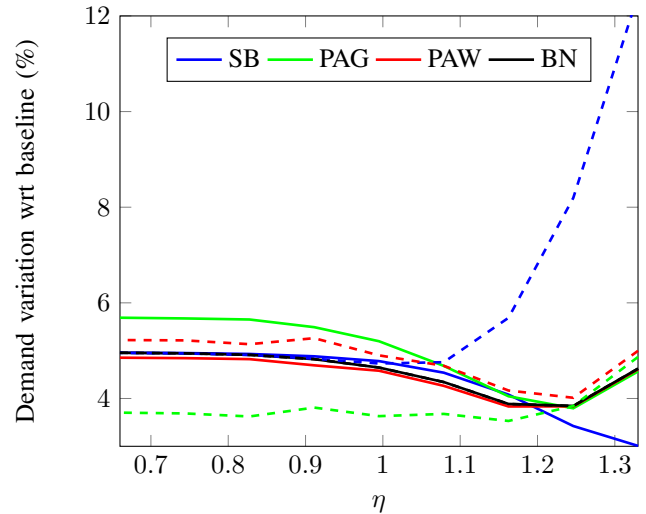
Power imbalances of Fig. 6 are also explained by the incorrect estimation of the flexibility provided by the distribution networks made by the different approaches. To illustrate this effect, we compute the relative difference between the approximate consumption of distribution networks $\widehat{P}_t^N$ and the baseline consumption, which is plotted in dashed lines in Fig. 7. The bold lines represent the relative difference between the actual consumption of distribution networks P_t^N and the same baseline demand. First, it can be observed that flexible customers allow for aggregate demand variations that range from 4% to 12%, on average. It can also be observed that PAG underestimates the flexibility provided by distribution networks for low congestion levels. Conversely, the SB approach overestimates the available flexibility of distributed networks when their operation is mainly driven by physical constraints. Finally, the proposed PAW approach is able to operate the transmission network with a very realistic estimation of the flexibility of the distribution networks, which is, in turn, very close to the actual flexibility levels determined by the centralized benchmark approach.

Similarly to Fig. 6, Fig. 8 plots the mean and the 5% and 95% percentiles of the social welfare loss with respect to the BN approach. Aligned with power imbalance results, the social welfare losses under the SB and PAG approaches are linked to high and low values of parameter η, respectively. More importantly, while the social welfare loss may reach values of 2% and 4% for the SB and PAG approaches, in that order, for some of the 100 hours analyzed, the PAW approach keeps this value below 0.1% for any network congestion level. That is, the proposed methodology to integrate transmission and distribution networks achieves the same social welfare as the BN for a wide range of power system conditions (described by the different demand and renewable capacity factors of the 100 hours) and network congestion of the distribution systems (modeled by parameter η). For completeness, Table II provides the sum of the social welfare for the 100 hours of the test set for some values of parameter η.
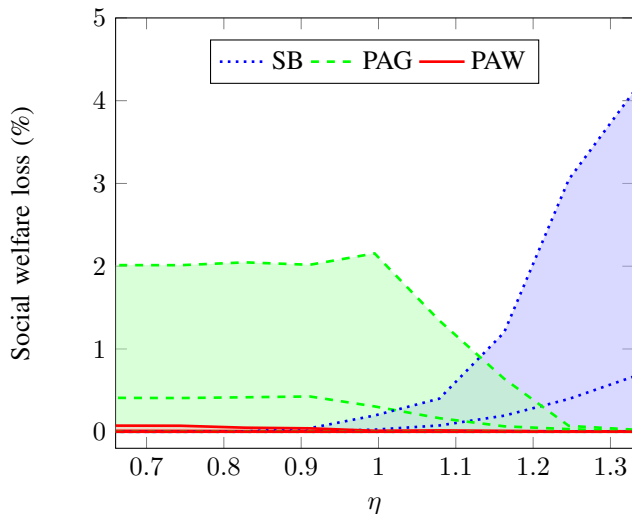
Fig. 8. Impact of distribution network congestion on social welfare.

TABLE II
Social welfare results in k€

η	BN	SB	PAG	PAW
0.66	3226.8	3226.8	3216.8	3226.3
1.00	3214.5	3213.5	3207.5	3214.4
1.33	2828.8	2814.1	2828.7	2828.7

It is also important to remark that social welfare losses in Fig. 8 are computed assuming that all generating units and consumers at the transmission network can react instantaneously to any real-time power imbalance without involving extra regulations costs. Therefore, these results are a lower bound of the actual social welfare losses that would happen in a more realistic setup in which flexibility resources are both limited and expensive.

Table III shows how the average relative social welfare loss (as illustrated in Fig. 8) is apportioned between the transmission and distribution systems, for various congestion levels η. Notably, the average loss in the SB and PAG cases disproportionally affects the transmission and distribution networks. Actually, there is a substantial net transfer of welfare from DSOs to the TSO. That is, the SB and PAG approaches delegate the bulk of the costs of dealing with distribution congestion to the distributed energy resources themselves, which certainly puts into question the ability of these methods to effectively integrate distribution into transmission operations. In contrast, the proposed PAW approach considerably mitigates this effect, or even reverses it, thus ensuring that distribution issues are also taken care of by transmission

TABLE III
Allocation of average social welfare loss (in percent with respect to BN) between transmission and distribution

	SB		PAG		PAW	
η	TSO	DSO	TSO	DSO	TSO	DSO
0.66	0.00%	0.00%	-4.98%	5.39%	0.46%	-0.45%
1.00	-1.38%	1.41%	-4.17%	4.47%	0.04%	-0.04%
1.33	-12.00%	12.67%	-0.54%	0.55%	-0.10%	0.10%

TABLE IV
Computational time results

	Min time (s)	Average time (s)	Max time (s)	Speedup
BN	0.88	1.66	15.27	-
SB	0.17	0.30	1.83	5.5x
PAG	0.02	0.03	0.20	55.3x
PAW	0.04	0.06	0.38	27.7x

resources.

Looking at Figures 6, 7 and 8, we can conclude that the effectiveness of the proposed PAW method with respect to other approaches depends on the network characteristics, and more particularly, on the congestion level of the distribution networks. Indeed, if the distribution networks never experience voltage congestion, SB outperforms PAW in terms of power imbalance and social welfare loss. On the other hand, for highly congested networks, PAG and PAW provide almost identical results. In conclusion, the use of the proposed PAW approach is more relevant for those systems in which the level of congestion of distribution networks vary significantly depending on the operating conditions.

Finally, Table IV compares the maximum, average and minimum computational times for the four approaches. The average speedup factor between each method and the benchmark is also provided in the last column. Due to the high number of variables and constraints of model (7), the SB's speedup factor is relatively low. In contrast, since PAG and PAW characterize the response of each distribution network through a constant value or a step-wise bidding curve, respectively, the computational savings are more substantial.

VI. CONCLUSION

Motivated by the proliferation of distributed energy resources, new TSO-DSO coordination strategies are required to take full advantage of these resources in the operation of the transmission system. Existing decomposition/decentralized methods are able to yield the same operating decisions as centralized benchmarks at lower computational costs. However, these approaches require access to proprietary information or fail-safe real-time communication infrastructures. Alternatively, our approach only uses offline historical data at substations to learn the price response of the distribution networks in the form of a non-increasing bidding curve that can be easily embedded into current procedures for transmission operations. In addition, this data set can be enriched with some covariates that have predictive power on the response of the distribution networks.

We have benchmarked our approach against an idealistic model that fully centralizes the coordination of distribution and transmission operations. We have also compared it with other approximate approaches that either ignore the technical constraints of the distribution networks or the price-sensitivity of DERs. The conducted numerical experiments reveal that our approach systematically delivers small differences with respect to the fully centralized benchmark in terms of power imbalances and social welfare regardless of the level of congestion of the distribution grids. In return, our approach is

computationally affordable and consistent with current market practices, and allows for decentralization.

Future work should be directed to assessing whether these results remain valid, and to which extent, for meshed distribution networks and DERs with more complex price responses, e.g. thermostatically controlled loads. Furthermore, in this research, we have only considered contextual information pertaining to continuous random variables (electricity demand, renewable power generation, etc.). Therefore, a relevant avenue for future research is to extend the proposed approach to work with binary variables too, such as those describing the network topology or the in-service/out-of-service status of some network components.

REFERENCES

[1] Z. Li, Q. Guo, H. Sun, and J. Wang, "Coordinated economic dispatch of coupled transmission and distribution systems using heterogeneous decomposition," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4817–4830, 11 2016.

[2] ——, "Coordinated transmission and distribution AC optimal power flow," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 1228–1240, 2018.

[3] G. Migliavacca, M. Rossi, D. Six, M. Džamarija, S. Horsmanheimo, C. Madina, I. Kockar, and J. M. Morales, "Smartnet: H2020 project analysing TSO–DSO interaction to enable ancillary services provision from distribution networks," *CIRED-Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 1998–2002, 2017.

[4] G. De Zotti, S. A. Pourmousavi, J. M. Morales, H. Madsen, and N. K. Poulsen, "A control-based method to meet TSO and DSO ancillary services needs by flexible end-users," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1868–1880, 2019.

[5] Z. Li, Q. Guo, H. Sun, and J. Wang, "Impact of coupled transmission-distribution on static voltage stability assessment," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3311–3312, 2017.

[6] G. Valverde, D. Shchetinin, and G. Hug-Glanzmann, "Coordination of distributed reactive power sources for voltage support of transmission networks," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1544–1553, 2019.

[7] Z. Li, Q. Guo, H. Sun, and J. Wang, "A new lmp-sensitivity-based heterogeneous decomposition for transmission and distribution coordinated economic dispatch," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 931–941, 2018.

[8] H. Gerard, E. I. Rivero Puente, and D. Six, "Coordination between transmission and distribution system operators in the electricity sector: A conceptual framework," *Utilities Policy*, vol. 50, pp. 40–48, 2018.

[9] A. Kargarian, J. Mohammadi, J. Guo, S. Chakrabarti, M. Barati, G. Hug, S. Kar, and R. Baldick, "Toward distributed/decentralized dc optimal power flow implementation in future electric power systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2574–2594, 2018.

[10] Z. Yuan and M. R. Hesamzadeh, "Hierarchical coordination of TSO-DSO economic dispatch considering large-scale integration of distributed energy resources," *Appl. Energy*, vol. 195, pp. 600 – 615, 2017.

[11] M. Bragin, Y. Dvorkin, and A. Darvishi, "Toward coordinated transmission and distribution operations," in *2018 IEEE Power Energy Society General Meeting (PESGM)*, 2018, pp. 1–5.

[12] A. Nawaz, H. Wang, Q. Wu, and M. Kumar Ochani, "TSO and DSO with large-scale distributed energy resources: A security constrained unit commitment coordinated solution," *Int. Trans. Electr. Energy Syst.*, vol. 30, no. 3, p. e12233, 2020, e12233 ITEES-19-0196.R3.

[13] A. Mohammadi, M. Mehrtash, and A. Kargarian, "Diagonal quadratic approximation for decentralized collaborative tso+dso optimal power flow," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2358–2370, 2019.

[14] T. Ding, C. Li, C. Huang, Y. Yang, F. Li, and F. Blaabjerg, "A hierarchical modeling for reactive power optimization with joint transmission and distribution networks by curve fitting," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2739–2748, 9 2018.

[15] Z. Li, T. Xu, Q. Guo, H. Sun, and J. Wang, "A response-function-based coordination method for transmission - distribution-coupled ac opf," in *2018 IEEE/PES Transmission and Distribution Conference and Exposition (T D)*, 2018, pp. 1–5.

[16] R. Mieth and Y. Dvorkin, "Online learning for network constrained demand response pricing in distribution systems," *IEEE Trans. Smart Grid*, vol. 11, pp. 2563–2575, 5 2020.

[17] A. Nawaz and H. Wang, "Stochastically coordinated transmission and distribution system operation with large-scale wind farms," *CSEE Journal of Power and Energy Systems*, pp. 1–10, 2020.

[18] D. Biesntock, *Electrical transmission system cascades and vulnerability*, 2015.

[19] A. Papavasiliou, "Analysis of distribution locational marginal prices," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4872–4882, 9 2018.

[20] R. Mieth and Y. Dvorkin, "Data-driven distributionally robust optimal power flow for distribution systems," *IEEE Control Systems Letters*, vol. 2, no. 3, pp. 363–368, 7 2018.

[21] A. Hassan, Y. Dvorkin, D. Deka, and M. Chertkov, "Chance-constrained ADMM approach for decentralized control of distributed energy resources," in *20th Power Systems Computation Conference, PSCC 2018*. Institute of Electrical and Electronics Engineers Inc., 8 2018.

[22] H. Le Cadre, I. Mezghani, and A. Papavasiliou, "A game-theoretic analysis of transmission-distribution system operator coordination," *Eur. J. Oper. Res.*, vol. 274, no. 1, pp. 317–339, 4 2019.

[23] J. Yu, Y. Weng, and R. Rajagopal, "PaToPa: A data-driven parameter and topology joint estimation framework in distribution grids," *IEEE Trans. Power Syst.*, vol. 33, no. 4, pp. 4335–4347, 7 2018.

[24] S. Chen, J. Ping, X. Le, Z. Yan, X. Xu, L. Yao, and J. Xi, "Forming bidding curves for a distribution system operator," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5389–5400, 9 2018.

[25] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, ser. Springer Series in Statistics. Springer New York, 2009.

[26] "OMIE," 2021, available (online): https://www.omie.es/es.

[27] V. Bucarey, M. Labbé, J. M. Morales, and S. Pineda, "A dynamic programming approach to segmented isotonic regression," *Omega*, vol. 105, p. 102516, 2021.

[28] I. Peña, C. B. Martinez-Anido, and B. Hodge, "An extended IEEE 118-bus test system with high renewable penetration," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 281–289, 2018.

[29] A. Hassan, R. Mieth, M. Chertkov, D. Deka, and Y. Dvorkin, "Optimal load ensemble control in chance-constrained optimal power flow," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5186–5195, 2018.

[30] A. Hassan and Y. Dvorkin, "Energy storage siting and sizing in coordinated distribution and transmission systems," *IEEE Trans. Sustain. Energy*, vol. 9, no. 4, pp. 1692–1701, 10 2018.

[31] OASYS, "Data 118TN+33DN," *GitHub repository (https://github.com/groupoasys/118TN33DN)*, 2021.