Distributed-MPC with Data-Driven Estimation of Bus Admittance Matrix in Voltage Stabilization

Ramij R. Hossain, Student Member, IEEE, Ratnesh Kumar, Fellow, IEEE

Abstract—This paper presents a distributed model-predictive control (MPC) design for real-time voltage stabilization in power systems, allowing the bus admittance matrix $Y = G + jB$ to be not known a priori (it may be time-varying), and so is estimated online. The prevalent control designs are either centralized and optimal but lack scalability and are subject to network attacks, or decentralized that are scalable and less vulnerable to attacks but are suboptimal. The proposed distributed solution offers the attractive features of both types of schemes, namely, optimality, scalability, as well as enhanced security to network attacks. In addition, since acquiring the exact knowledge of the line conductance and susceptance, which are required to form $Y$, is in general challenging, the presented framework integrates data-driven estimation of $Y$ to circumvent this challenge. We first introduce the centralized version of the formulation, and next transfer it to a distributed version for efficiency, scalability, and attack-resilience, leveraging the graph structure of the power system. Only local computation and communication are used for (i) computing local control via distributed optimization (which is solved by the alternating direction method of multipliers ADMM), as well as (ii) data-driven estimation of $Y$. The performance of the proposed methodology is validated using numerical examples of the IEEE-30 Bus and IEEE-57 Bus systems.

Index Terms—Distributed Optimization, MPC, ADMM, Voltage Control, Data-Driven method

I. INTRODUCTION

Traditionally the optimization problems that emerge in power system operations have been solved using model-based centralized algorithms. Voltage control, which is one of the fundamental problems in power systems operation, is no exception [1], [2]. With the increasing penetration of renewable energy resources, dynamic loads, energy storage devices, and plug-in electric vehicles, the power system operation is currently experiencing a paradigm shift. In this context, it is important to note that designing real-time system-wide control, commonly referred to as wide-area control [3], [4] is of utmost importance for the resiliency and security of large-scale power systems. While centralized controllers are suitable for system-wide optimization; there are specific challenges in implementing centralized control frameworks for real-time system operation. A centralized implementation needs the utilities to collect and communicate PMU measurements to a central location to compute the necessary controls via centralized optimization. A centralized framework thus requires robust communication infrastructure and computation resources capable of handling high data volume. Plus, this framework becomes vulnerable to cyber-attacks due to the dependence on a single centralized solver. Therefore, there is a need for a distributed alternative yet ensuring optimality.

This is additionally reinforced by the fact that, both in the case of transmission and distribution systems, the structure of the operation is becoming more distributed [5], which, in turn, necessitates the incorporation of decentralized controls using local measurements and computations. For example, sometimes voltage control problems are solved utilizing droop-based local reactive compensations, [6], [7]. Similarly, some gradient-based techniques can also be implemented in a complete decentralized fashion [8], [9]. However, an ad-hoc decentralization without considering the impact of inter-area connections in power systems can have an adverse effect on overall system performance [10], [11], and that is why a systematic distributed optimization in power systems is gaining attention in recent years [12]–[15] for efficiency as well as cyber-security of computation and control.

An additional issue in model-based design is the requirement of a sound knowledge of the system parameters, e.g., the knowledge of bus admittance matrix $Y$. Traditionally, the elements of $Y$ are populated using the values obtained during the system design. But these values may not be readily updateable while the system is operating, upon any modifications of the network topology and device parameters [16]. To this end, designing control using data-driven estimation of $Y$ constitutes a vital problem in the real-time operation of power systems, and further, for the reasons mentioned earlier, that control must also be distributed.

These serve as the motivation for our work on data-driven distributed control in model-predictive (D3-MPC) framework as presented here. A scalable distributed model-predictive algorithm for wide-area voltage control of a power system with guaranteed optimality, while involving communication limited to among the neighbors and when the bus admittance matrix is not known a priori or may be time-varying (and so needs to be estimated online), has not been developed yet. This paper fills this gap by proposing a predictive-control-based data-driven distributed optimization algorithm for parameter estimation and voltage control in a wide-area setting.

A. Related works

The existing studies on integrating local communication-based distributed computation in voltage control have mostly explored either feedback-based voltage regulation or an optimal power/reactive power flow (OPF/ORPF) analysis. Plus,
distributed optimization in power system applications is predominantly done for distribution network [14], using convex relaxations (e.g., SDP and SOCP relaxations) or model approximations (e.g., the Linearized DistFlow (LinDistflow) or DC power flow approximation). [17], [18] cast the voltage control as an ORPF problem and utilized the DC power flow equations to model a distribution system. A linearly-constrained quadratic optimization problem is solved in a distributed fashion in [17], [18]. An ADMM based distributed solution to the optimal reactive control problem can be found in [19]. This work considered a LinDistflow approximation of a radial distribution system. The authors in [20] proposed an ADMM approach to finding an optimal setting of reactive power resources in a balanced radial network leveraging Distflow approximation (a variant of branch flow model BFM), and later extended the approach to tackling unbalanced radial meshed networks.

[21] used bus injection model (BIM) combined with SDP-based convex relaxation to design the distributed controllers for voltage regulation. The proposed distributed algorithm is only efficient for solving convexified SDPs on acyclic networks. [22] proposed an ADMM based distributed solver to maintain the voltage profile close to a given reference value in an unbalanced distribution network. This work utilized a sparse sensitivity-based linearization around a flat voltage profile (linear q-v model) to avoid the computational challenges imposed by the nonlinear couplings of system equations. A feedback-based online correction was needed to tackle the approximations used in the optimization problem. ADMM has also been used in power system applications in [23]. A distributed feedback controller using the dual-ascent method was proposed in [24] with a combined objective of maximizing the distance of voltage collapse and voltage regulation. This approach also relied on the convex reformulation of the optimization problem by considering the linearized system approximation. The Optimal Distributed Feedback Voltage Control (OPTDIST-VC) proposed in [25] is again based on the LinDistflow approximation of the power system and utilizes the network sparsity to design the primal-dual based distributed control technique.

The distributed voltage control works mentioned above are (i) heavily reliant on the system model and its approximation; most of the previous works utilized the Distflow/LinDistflow approximated for the radial distribution system, and (ii) mostly based on the convexified OPF/ORPF solution. The convex relaxation of the nonconvex OPF/ORPF problem does not guarantee the feasibility of the obtained solution, requiring those to be verified separately [14].

B. Our Contributions

In contrast to these earlier works listed above, we introduce a distributed voltage control framework in the model predictive control (MPC) setting, considering the generic and accurate power system modeling (i.e., without any model approximations as above). MPC computes the control variables iteratively at each control instant. Using a linearization around a predicted trajectory, the optimization problem is posed with respect to incremental changes of control variables. Our distributed MPC-based formulation decomposes the optimization among local control agents, relying on communications from only their local neighbors. We employ ADMM to solve the MPC optimization in a distributed manner. Due to the convexity, the ADMM based distributed solution also guarantees asymptotic convergence to globally optimal solution [26], [27].

Further, in contrast to the earlier works in distributed voltage control that need the knowledge of the system parameters, specifically line resistances and reactances to form $\mathbf{Y}$, our work allows unknown line resistances and reactances: We propose a method for data-driven estimation of the bus admittance matrix $\mathbf{Y}$ using the real-time SCADA/PMU data and is also performed distributively involving local communication among neighboring agents.

Summarily, we have developed and validated a data-driven distributed predictive control formulation that does not require the knowledge of system parameters, rather estimates those distributively in real-time. It offers optimality, scalability, and enhanced security to network attacks (in contrast to prevalent centralized schemes that are optimal but not scalable or attack-resilient or decentralized schemes that are scalable and more attack-resilient but suboptimal). The proposed methodology is tested against IEEE 30 Bus and 57 Bus systems to (a) achieve the desired voltage performance using data-driven estimation of $\mathbf{Y}$, (b) satisfy the network constraints, and (c) maintain the reactive power constraints throughout, while only utilizing the local measurements and communications.

C. Organization

The rest of the paper is organized as follows: Section II describes the power system model and MPC-based centralized formulation of the optimization problem for voltage stabilization. Section III introduces the proposed ADMM-based distributed solution of the optimization problem, while the data-driven estimation of system parameters is presented on IV. Results for the case studies are provided in Section V. Finally, Section VI concludes the work and the future directions.

II. POWER SYSTEM MODEL AND PROBLEM FORMULATION

A power system connectivity of buses via the lines can be represented by an undirected network graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$, where the set of buses includes $\mathcal{N} = \{0, \cdots, N\}$ and $\mathcal{E}$ represents the set of lines connecting the buses, with line $l_{ik} \in \mathcal{E}$ connecting buses $i, k \in \mathcal{N}$. The magnitude and angle of voltage at bus $i$ are denoted $V_i$ and $\theta_i$ respectively. We let bus 0 to be the slack/reference bus, so by convention, $V_0 = 1$ p.u. and $\theta_0 = 0^\circ$. The symmetric bus admittance matrix $\mathbf{Y} = [y_{ik}] \in \mathbb{C}^{[\mathcal{N}] \times [\mathcal{N}]}$ of the given network is defined as:

\[
y_{ik} = \begin{cases} y_i + \sum_{k=0, k \neq i}^N y_{ik}, & \text{if } i = k \\ -y_{ik}, & \text{otherwise} \end{cases}, \tag{1}
\]

where it holds that $y_{ik} = y_{ki} \neq 0$ if $l_{ik} \in \mathcal{E}$ (i.e., buses $i$ and $k$ are connected) and otherwise $y_{ik} = y_{ki} = 0$, and $y_i$ denotes the admittance to the ground at bus $i$. We write the
complex admittance as, \( Y_{ik} = G_{ik} + jB_{ik} \), with \( G_{ik} \) := real part (conductance) and \( B_{ik} \) := imaginary part (susceptance).

The active and reactive power generations at any bus \( i \) are denoted by \( P^G_i, Q^Q_i \), respectively, while \( P^D_i, Q^D_i \) represent the active and reactive power demands. Therefore the net injections of active and reactive powers at bus \( i \) are given by:

\[
P^{in}_i = P^G_i - P^D_i, \quad Q^{in}_i = Q^Q_i - Q^D_i.
\]

From power flow relation at bus \( i \) we have:

\[
P^{in}_i = \sum_{k=0}^{N} V_i V_k \left[ G_{ik} \cos (\theta_i - \theta_k) + B_{ik} \sin (\theta_i - \theta_k) \right] \tag{2a}
\]

\[
Q^{in}_i = \sum_{k=0}^{N} V_i V_k \left[ G_{ik} \sin (\theta_i - \theta_k) - B_{ik} \cos (\theta_i - \theta_k) \right] \tag{2b}
\]

It is important to note that the power flow equations of bus \( i \) given in (2) depend only on the variables of the bus itself and of its neighboring buses. This follows from the graph structure of the network, which implies that \( G_{ik} = B_{ik} = 0 \) whenever the buses \( i \) and \( k \) are not connected by a line. Letting \( N_i := \{k : i \in E\} \) denote the “neighboring buses” of bus \( i \), the power flow equations (2) can be represented as follows:

\[
P^{in}_i = g^P_i(V_i, \theta_i, V_k, \theta_k \mid k \in N_i), \tag{3a}
\]

\[
Q^{in}_i = g^Q_i(V_i, \theta_i, V_k, \theta_k \mid k \in N_i). \tag{3b}
\]

With the proliferation of renewable energy generation and the variable nature of the loads, maintaining voltage trajectories close to the desired reference value of \( V_{ref} = 1.00 \text{ p.u.} \) is challenging, requiring a well-defined control design. Typical profiles of utility level renewable generation are collected from [28], and shown in Fig. 1 for three different days.

However, with the advent of AI-based prediction techniques, nowadays, an accurate prediction of up to an hour-ahead generation and load profiles has become possible for most utilities and system operators [29], [30]. The prediction error is typically bounded to 3-5% in the data collected at [28]. Accordingly in our work, along with current information of system variables, we also employ the predicted generation and load profiles already available at the utilities through the use of the well-established AI-based forecasting methods:

At any control time instant \( t \), the current bus voltages and angles \([V^*_t, \theta^*_t]\) as well as the current active and reactive power injections \([P_{i,t}, Q_{i,t}]\) are known for all \( i \in N \), and additionally, the predicted generations \([P_{i,t+1}^{G,t+1}, Q_{i,t+1}^{Q,t+1}]\) and predicted demand profiles \([P_{i,t+1}^{D,t+1}, Q_{i,t+1}^{D,t+1}]\), and thereby the predicted injections \([P_{i,t+1}^{in,t+1}, Q_{i,t+1}^{in,t+1}]\) for the next control time instant \( t + 1 \) are also available from the forecast data.

Under the knowledge of the above data at control instant \( t \), we next present a predictive control formulation to compute an optimal reactive power correction at the designated buses for time instant \( t + 1 \). It turns out that the existing predictive methods employ sensitivity-based formulation [1], [31], in which the computation of sensitivity of voltages to controls is required, and which in turn, requires system-wide communication, and as such is not amenable to a distributed control framework. Hence, here we present a novel centralized predictive control formulation that does not employ control sensitivity computation, and as a result permits its distributed implementation. The centralized formulation is presented first, followed by its distributed implemented in the next section.

For predicted injections of active and reactive powers \([P_{i,t+1}^{in,t+1}, Q_{i,t+1}^{in,t+1}]\) and an unknown reactive compensation of \([u_{i+1}^{t+1}]\), the power flow equations (3) at bus \( i \) for time \( t + 1 \) become:

\[
g^P_i(V^*_t + \Delta V^t + \Delta V_{i,t+1}^t, \theta^*_t + \Delta \theta^t + \Delta \theta_{i,t+1}^t \mid k \in N_i) = \hat{P}^{in,t+1}_i,
\]

\[
g^Q_i(V^*_t + \Delta V^t + \Delta V_{i,t+1}^t, \theta^*_t + \Delta \theta^t + \Delta \theta_{i,t+1}^t \mid k \in N_i) = Q^{in,t+1}_i + u_{i+1}^{t+1}, \tag{4a, b}
\]

where \([\Delta V^t, \Delta \theta^t] \mid k \in N_i \) are unknown predicted values for bus \( i \) at \( t + 1 \). (Similarly, \([\Delta V_k^t, \Delta \theta_k^t] \mid k \in N_i \) are unknown predicted values for all neighboring buses \( k \in N_i \).) We can analogously express the known predicted changes in injected active and reactive powers as:

\[
\Delta P^{in,t+1}_i = \hat{P}^{in,t+1}_i - P^{in,t}_i, \quad \Delta Q^{in,t+1}_i = \hat{Q}^{in,t+1}_i - Q^{in,t}_i.
\]

Using Taylor series approximation on (4), we obtain:

\[
\left( \frac{\partial g^P_i}{\partial V_i} \right)^t \Delta V^t_{i,t+1} + \left( \frac{\partial g^P_i}{\partial \theta_i} \right)^t \Delta \theta^t_{i,t+1} + \sum_{k \in N_i} \left[ \left( \frac{\partial g^P_i}{\partial V_k} \right)^t \Delta V^t_k + \left( \frac{\partial g^P_i}{\partial \theta_k} \right)^t \Delta \theta^t_k \right] = \Delta P^{in,t+1}_i, \tag{5a}
\]

\[
\left( \frac{\partial g^Q_i}{\partial V_i} \right)^t \Delta V^t_{i,t+1} + \left( \frac{\partial g^Q_i}{\partial \theta_i} \right)^t \Delta \theta^t_{i,t+1} + \sum_{k \in N_i} \left[ \left( \frac{\partial g^Q_i}{\partial V_k} \right)^t \Delta V^t_k + \left( \frac{\partial g^Q_i}{\partial \theta_k} \right)^t \Delta \theta^t_k \right] = \Delta Q^{in,t+1}_i + u_{i+1}^{t+1}, \tag{5b}
\]

in which \([\Delta P^{in,t+1}_i, \Delta Q^{in,t+1}_i] \) are known predicted values, while \([\Delta V^t_i, \Delta \theta^t_i, u_{i+1}^{t+1}] \) are unknown but get determined through the following proposed Centralized MPC (C-MPC):

\[
\min_{\Delta V^t_{i,t+1}, \Delta \theta^t_{i,t+1}, u_{i+1}^{t+1}} \sum_{i=1}^N \left[ ||V^t_i + \Delta V^t_{i,t+1} - V^t_{ref}||_2 + ||u_{i+1}^{t+1}||_2 \right]
\]

(6a)
subject to, \( \forall i \in \mathbb{N} \):  

\[
\left( \frac{\partial g^P_i}{\partial \theta_i} \right)^T \Delta V_i^{t+1} + \left( \frac{\partial g^Q_i}{\partial \theta_i} \right)^T \Delta \theta_i^{t+1} + \sum_{k \in \mathbb{N}_i} \left[ \left( \frac{\partial g^P_i}{\partial \theta_k} \right)^T \Delta V_k^{t+1} \right] = \Delta P_i^{in,t+1}, \quad (6b)
\]

\[
\left( \frac{\partial g^Q_i}{\partial \theta_i} \right)^T \Delta V_i^{t+1} + \left( \frac{\partial g^Q_i}{\partial \theta_i} \right)^T \Delta \theta_i^{t+1} + \sum_{k \in \mathbb{N}_i} \left[ \left( \frac{\partial g^Q_i}{\partial \theta_k} \right)^T \Delta V_k^{t+1} \right] = \Delta Q_i^{in,t+1} + u_i^{t+1}, \quad (6c)
\]

\[
u_i^{t+1} = \left. \frac{\partial g^P_i}{\partial \theta_i} \right|_{\theta_i = \theta_i} \Delta V_i^{t+1} \leq \Delta \theta_i^{t+1} \leq \left. \frac{\partial g^P_i}{\partial \theta_i} \right|_{\theta_i = \theta_i} \Delta V_i^{t+1}, \quad (6d)
\]

\[
V_i^{\text{min}} \leq V_i^t + \Delta V_i^{t+1} \leq V_i^{\text{max}}, \quad (6e)
\]

\[
\Delta \theta_i^{\text{min}} \leq \Delta \theta_i^{t+1} \leq \Delta \theta_i^{\text{max}}. \quad (6f)
\]

In the C-MPC formulation of (5), the optimization variables are defined as:  
\( \Delta V^{t+1} := [\Delta V_1^{t+1}, \ldots, \Delta V_{N-1}^{t+1}]^T \), \( \Delta \theta^{t+1} := [\Delta \theta_1^{t+1}, \ldots, \Delta \theta_N^{t+1}]^T \), and \( u_i^{t+1} := [u_1^{t+1}, \ldots, u_{N-1}^{t+1}]^T \). While \( \{u_i^{t+1}\} \) locally impacts the model equation of respective bus \( i \). Keeping this in mind, we separate the optimization variables in (6) into two categories: a) public variables: \( \Delta V^{t+1}, \Delta \theta^{t+1} \) and b) private variables: \( u_i^{t+1} \). To be able to reformulate (6) as a distributed optimization, we next introduce at each bus a local replica copy of each public variable appearing its model equation, namely, at each bus \( i \in \mathbb{N} \), we introduce a vector

\[
x_i := \left( \frac{\Delta V^{t+1} \Delta \theta^{t+1}}{\Delta \theta_i^{t+1}} \right)_{k \in \{i\} \cup \mathbb{N}_i}^T \in \mathbb{R}^{2 \times (|\mathbb{N}_i|+1)}.
\]

In addition to \( x_i \), each bus \( i \) maintains a separate copy of its own local variables as the true values of the respective variables:

\[
z_i := \left( \Delta V_i^{t+1} \Delta \theta_i^{t+1} \right) \in \mathbb{R}^2.
\]

The true variables in \( z \) \( \in \mathbb{R}^{2 \times (|\mathbb{N}_i|+1)} \) and the local replica variables in \( x := [x_i]_{i \in \mathbb{N}}^T \) must together satisfy the self consistency constraints expressed as follows:

\[
\forall i \in \mathbb{N} : x_i = E_i z, \quad (7a)
\]

\[
\forall k \in \{i\} \cup \mathbb{N}_i, l \in \mathbb{N} : E_i(k,l) := \begin{cases} I_{2 \times 2} =: E_{ik} & \text{if } x_{i(k)} = z_l \\ 0_{2 \times 2} & \text{otherwise} \end{cases} \quad (7b)
\]

Note \( E_i \in \mathbb{R}^{2 \times (|\mathbb{N}_i|+1) \times 2 \times |\mathbb{N}|} \), and if we let \( E_i(k) \) denotes its \( k \)-th column \( k \in \mathbb{N} \), (7a) can be written as (13b).

**Illustrative Example:** The distributed computation framework involving the new set of variables defined above is illustrated using a 4-Bus system shown in Fig. 2A. Here the bus 0 is the slack bus, therefore \( V_0 = 1 \) p.u. and \( \theta_0 = 0^p \). These remain fixed, therefore \( \Delta V_0 = 0 \), and \( \Delta \theta_0 = 0 \). The graph representation for the rest of the network is shown in Fig. 2B.
The details of the local copies and true public variables are given next.

| Bus No | Neighbors | Public Local Copy | Public True | Private |
|--------|-----------|-------------------|-------------|---------|
| Bus-1  | $N_1 = \{2\}$ | $x_1 = [x_{1,1}, x_{2,1}]^T$ | $z_1$ | $u_1$ |
| Bus-2  | $N_2 = \{1, 3\}$ | $x_2 = [x_{2,2}, x_{2,3}, x_{3,2}]^T$ | $z_2$ | $u_2$ |
| Bus-3  | $N_3 = \{2\}$ | $x_3 = [x_{3,3}, x_{3,2}]^T$ | $z_3$ | $u_3$ |

@Bus-1: \[
\begin{align*}
(x_{1,1}) &= (\Delta V_1^{t+1}, \Delta \theta_1^{t+1}), \\
(x_{1,2}) &= (\Delta V_2^{t+1}, \Delta \theta_2^{t+1}).
\end{align*}
\]

@Bus-2: \[
\begin{align*}
(x_{2,2}) &= (\Delta V_2^{t+1}, \Delta \theta_2^{t+1}), \\
(x_{2,1}) &= (\Delta V_1^{t+1}, \Delta \theta_1^{t+1}), \\
(x_{2,3}) &= (\Delta V_3^{t+1}, \Delta \theta_3^{t+1}).
\end{align*}
\]

@Bus-3: \[
\begin{align*}
(x_{3,3}) &= (\Delta V_3^{t+1}, \Delta \theta_3^{t+1}), \\
(x_{3,2}) &= (\Delta V_2^{t+1}, \Delta \theta_2^{t+1}).
\end{align*}
\]

For the self consistency of the variables we need: \[
\begin{align*}
x_{1,1} &= x_{2,1} = z_1, \\
x_{2,2} &= x_{1,2} = x_{3,2} = z_2, \\
x_{3,3} &= x_{2,3} = z_3.
\end{align*}
\]

Hence we can define $E_i^\prime$ s as follows:

\[
E_1 = \begin{bmatrix}
E_{11} & 0 & 0 \\
0 & E_{12} & 0 \\
0 & 0 & I
\end{bmatrix}
\]

\[
E_2 = \begin{bmatrix}
0 & E_{22} & 0 \\
E_{21} & 0 & 0 \\
0 & 0 & E_{23}
\end{bmatrix}
\]

\[
E_3 = \begin{bmatrix}
0 & 0 & E_{33} \\
0 & E_{32} & 0 \\
0 & 0 & I
\end{bmatrix}
\]

The augmented lagrangian for (13) can be written as in (14), where $\lambda$ is the lagrangian multiplier, and $\rho$ is a penalty constant:

\[
L_\rho^i(x_i, u_i, z_i, \lambda_i) = J_i(x_i, u_i) + \lambda_i^T(x_i - \sum_{k \in \{i\} \cup N_i} E_i(k) z_k) + \frac{\rho}{2} \left\| x_i - \sum_{k \in \{i\} \cup N_i} E_i(k) z_k \right\|^2.
\]

Then following [27], the dual problem of (13) can be solved by Algorithm 1, which then also yields the centralized optimum of (6) as explained below in Remark 1.

**Algorithm 1 Alternating Direction Method of Multiplier (ADMM)**

1. Parallexly $\forall i \in N$:
   - Initialize $\lambda_i = z_i = 0$.
   - repeat
     1. Update $(x_i^+, u_i^+) = \text{argmin}_{(x_i, u_i) \in C_i} L_\rho^i(x_i, u_i, z_i, \lambda_i)$
     2. Communicate $x_i^+$ to all neighbor $k \in N_i$
     3. Update $z_i^+ = \frac{1}{|N_i + 1|} \sum_{k \in \{i\} \cup N_i} E_k(i)^T(x_k + \frac{1}{\rho} \lambda_k)$
     4. Communicate $z_i^+$ to all neighbors $k \in N_i$
     5. Update $\lambda_i^+ = \lambda_i + \rho(x_i^+ - \sum_{k \in \{i\} \cup N_i} E_i(k) z_k)$
     6. Communicate $\lambda_i^+$ to all neighbors $k \in N_i$
   - until convergence

**Remark 1:** Following [27], whenever the local objective functions $J_i(x_i, u_i), \forall i \in N$ are closed, proper, and convex and the overall augmented Lagrangian,

\[
L_\rho(x, u, z, \lambda) = \sum_{i=1}^N L_\rho^i(x_i, u_i, z_i, \lambda_i),
\]

has a saddle point, Algorithm 1 has the property that the residuals $\|x_i - E_i z_i\|$ converge asymptotically to zero for all $i \in N$ and value of $J(x, u) = \sum_{i=1}^N J_i(x_i, u_i)$ converges asymptotically to the primal optimum. In our case, (a) the C-MPC problem (6) is convex, and hence its overall augmented
IV. DATA-DRIVEN ESTIMATION OF ADMITTANCE MATRIX

The above distributed optimization framework requires the computation of the required partial derivatives in (10) (that appear in (9) or equivalently in (6b)-(6c)). The partial derivatives are of the functions \( g_i^p (\cdot) \) and \( g_i^q (\cdot) \) introduced in (3), and their values are given as follows:

\[
g_i^p (V_i, \theta_i, V_k, \theta_k) = V_i^2 G_{ii} + \sum_{k \in \mathbb{N}_i} V_i V_k \left[ G_{ik} \cos (\theta_i - \theta_k) + B_{ik} \sin (\theta_i - \theta_k) \right], \quad (16a)
\]

\[
g_i^q (V_i, \theta_i, V_k, \theta_k) = -V_i^2 B_{ii} + \sum_{k \in \mathbb{N}_i} V_i V_k \left[ G_{ik} \sin (\theta_i - \theta_k) - B_{ik} \cos (\theta_i - \theta_k) \right]. \quad (16b)
\]

Here \( g_i^p (\cdot) \) and \( g_i^q (\cdot) \) depend on the entries of the admittance matrix \( Y \), for which from (1), \( Y_{ii} = y_i + \sum_{k=0, k \neq i}^N y_{ik} \equiv G_{ii} + j B_{ii} \). Hence we have:

\[
G_{ii} = g_i + \sum_{k=0, k \neq i}^N g_{ik} = g_i - \sum_{k \in \mathbb{N}_i} G_{ik}, \quad (17a)
\]

\[
B_{ii} = b_i + \sum_{k=0, k \neq i}^N b_{ik} = b_i - \sum_{k \in \mathbb{N}_i} B_{ik}. \quad (17b)
\]

Note in general, \( g_i = 0 \), and since the control goal is to compute the reactive compensations, we can also treat \( b_i = 0 \) by letting it appear in the form of the added control input \( u_i \). With this convention, we have: \( G_{ii} = -\sum_{k \in \mathbb{N}_i} G_{ik} \) and \( B_{ii} = -\sum_{k \in \mathbb{N}_i} B_{ik} \). Consequently, (16) can be simplified as:

\[
g_i^p (V_i, \theta_i, V_k, \theta_k) = \sum_{k \in \mathbb{N}_i} \left[ V_i V_k \cos (\theta_i - \theta_k) - V_i^2 \right] G_{ik}
\]

\[
+ V_i V_k \sin (\theta_i - \theta_k) B_{ik}, \quad (18a)
\]

\[
g_i^q (V_i, \theta_i, V_k, \theta_k) = \sum_{k \in \mathbb{N}_i} \left[ V_i V_k \sin (\theta_i - \theta_k) G_{ik} + V_i^2 \right]
\]

\[
+ \left\{ -V_i^2 - V_i V_k \cos (\theta_i - \theta_k) \right\} B_{ik}. \quad (18b)
\]

We can now compute the partial derivative terms, and assemble them in the matrix form (19). Clearly the values of the partial derivatives for bus \( i \in \mathbb{N} \) requires i) the current measurements of bus voltages and angles of buses \( \{i\} \cup \mathbb{N}_i \), and ii) the knowledge of \( G_{ik} \) and \( B_{ik} \), \( k \in \mathbb{N}_i \), which as shown next can be obtained from the power measurements. The standard expressions for \( P_{ik}, Q_{ik}, P_{ki}, Q_{ki} \) (the active and reactive power flows from bus \( i \) to \( k \) and the reverse) are as follows:

\[
P_{ik} = [V_i V_k \cos (\theta_i - \theta_k) - V_i^2] G_{ik} + V_i V_k \sin (\theta_i - \theta_k) B_{ik},
\]

\[
Q_{ik} = V_i V_k \sin (\theta_i - \theta_k) G_{ik} + [V_i^2 - V_i V_k \cos (\theta_i - \theta_k)] B_{ik},
\]

\[
P_{ki} = [V_i V_k \cos (\theta_k - \theta_i) - V_k^2] G_{ik} + V_k V_i \sin (\theta_k - \theta_i) B_{ik},
\]

\[
Q_{ki} = V_k V_i \sin (\theta_k - \theta_i) G_{ik} + [V_k^2 - V_k V_i \cos (\theta_k - \theta_i)] B_{ik},
\]

which we can note is of the form of \( A \Theta = b \), where:

\[
A = \begin{bmatrix}
V_i V_k \cos (\theta_i - \theta_k) - V_i^2 & V_i V_k \sin (\theta_i - \theta_k) \\
V_i V_k \sin (\theta_i - \theta_k) & V_i^2 - V_i V_k \cos (\theta_i - \theta_k) \\
V_k V_i \cos (\theta_k - \theta_i) - V_k^2 & V_k V_i \sin (\theta_k - \theta_i) \\
V_k V_i \sin (\theta_k - \theta_i) & V_k^2 - V_k V_i \cos (\theta_k - \theta_i)
\end{bmatrix}
\]

\[
b = \begin{bmatrix}
P_{ik} \\
P_{ki} \\
Q_{ik} \\
Q_{ki}
\end{bmatrix}, \quad \text{and} \quad \Theta = \begin{bmatrix}
G_{ik} \\
B_{ik}
\end{bmatrix}.
\]

Thus using the measurement data \( \{V_i, V_k, \theta_i, \theta_k, P_{ik}, Q_{ik}, P_{ki}, Q_{ki} \mid k \in \mathbb{N}_i \} \) at each bus \( i \in \mathbb{N}_i \), we can compute the least square estimate (LSE) of \( G_{ik} \) and \( B_{ik} \) by minimizing:

\[
\min_{\Theta} \|A \Theta - b\|.
\]

The well-known solution of the LSE optimization is given by:

\[
\hat{\Theta} = \left( A^T A \right)^{-1} A^T b. \quad (20)
\]

The above data-driven estimation of \( \{G_{ik}, B_{ik} \mid i \in \mathbb{N}, k \in \mathbb{N}_i \} \) removes the need of any a priori knowledge of system parameters (the line resistances and reactances) during the optimization of (13) using Algorithm 1. It simply relies on i) the real-time active and reactive power flow of the transmission lines that are available from SCADA measurements [33], and ii) the voltage magnitudes and angles that are available from PMU measurements. Thereby our proposed methodology is Data-Driven Distributed MPC (D3-MPC). The complete steps of the proposed D3-MPC are enumerated in Algorithm 2.

Algorithm 2 Data-Driven Distributed MPC (D3-MPC)

Parallely \( \forall i \in \mathbb{N}_i \), for each control instant \( t \in [0, T] \) do

1. Obtain latest measurement data \( \{V_i, \theta_i, P_{ik}, Q_{ik} \mid k \in \mathbb{N}_i \} \).

2. Communicate with neighboring buses/agents \( k \in \mathbb{N}_i \), and receive the data \( V_k, \theta_k, P_{ki}, Q_{ki} \).

3. Using the collected data, estimate \( \hat{G}_{ik} \) and \( \hat{B}_{ik} \) for each \( k \in \mathbb{N}_i \) employing (20).

4. Compute the required partial derivatives (19) using the estimated values \( \hat{G}_{ik} \) and \( \hat{B}_{ik} \), and the latest measurements of \( \{V_i, \theta_i, V_k, \theta_k \mid k \in \mathbb{N}_i \} \).

5. Collect the predicted generation \( P_{i,k+1}^G, Q_{i,k+1}^G \) and demand profiles \( P_{i,k+1}^D, Q_{i,k+1}^D \).

6. Solve the MPC problem (13) by Algorithm 1.

7. From the output of the optimization, obtain the reactive compensation \( u_i^{k+1} \) and implement it for the next time step \( t + 1 \).

end for

V. TEST RESULTS

Here we validate the capability of our proposed D3-MPC framework using two examples of standard IEEE systems by way of tracking a specified reference voltage \( V_{\text{ref}} \) in the
face of uncertain time-varying load and renewable generations. For comparison of the proposed D3-MPC, we first examine the performance of C-MPC (centralized MPC) by allowing it also the benefit of knowing the bus admittance matrix $Y$. In contrast, for the performance of the proposed D3-MPC (Data-Driven Distributed MPC), we utilize (a) the estimated values of the admittance matrix parameters using the measurement data, and (b) ADMM for distributed optimization with communication limited to among the neighbors as described in Algorithm 2. For simulating a power system, we relied on PYPOWER, and for optimizing C-MPC as well as D3-MPC, we utilized CVXPY.

The validation results are first demonstrated for IEEE 30 bus system and next for the IEEE 57 bus system. For our implementation, we utilized intel(R) Core(TM) i7-4790 CPU @ 3.60GHz processor with 16 GB RAM. For distributed optimization, we employed Ray framework [34] which supports task parallelism via Ray remote functions.

A. Test Case of IEEE 30 Bus System

We utilized the standard IEEE 30 Bus network with six generators located at buses 1, 2, 13, 22, 23, and 27. We include renewable generations (solar as well as wind) at these generator buses. Moreover, to introduce a realistic setting, we consider a mix of conventional and renewable generations with 50% coming from renewable sources during the peak of renewable generations. The loads and the renewable generations are time-varying, however as discussed in the paragraph following Fig. 1, the prediction of load and renewable generation profiles are available, say at an interval of 15 minutes. For our validation purposes, we obtained real-world utility-scale data from [28] and generated the representative plots for such predictions as shown in Fig. 3. Under these representative profiles, in the case of no control compensation, the voltage values of multiple buses drop below the recommended $V_{\text{min}} = 0.95$ p.u. as can be seen in Fig. 4, that has a minimum voltage of around 0.91 p.u.

![Fig. 3. Predicted renewable generation (A), and load profile (B)](image)

Here we utilize the predictive control framework—the centralized version C-MPC formulated in (6) that assumes the knowledge of the admittances versus the data-driven distributed alternative D3-MPC of (13) in which not only the control computations are distributed, the admittances are not known a priori and are rather estimated online based on SCADA/PMU data—to mitigate the voltage deviations from a given reference $V_{\text{ref}} = 1$ p.u., ensuring that the voltage trajectories always remain within the [0.95, 1.05] p.u., by way of reactive compensations at individual buses (except the slack) under the reactive power limits of [-0.05, 0.05] p.u.

Fig. 3 shows the real-world renewable generation and load prediction profiles. Allowing for the prediction errors that range from 3-5% in practice (see [28]), we first evaluate the performance of the C-MPC scheme for 3 different cases of prediction errors—±2%, ±5%, ±7%. The results are presented in Fig. 5 from which it can be seen that the C-MPC scheme can manage the control inputs to achieve the desired voltage performance under the realistic uncertainties of the predictions.

Next we examine the performance of D3-MPC. Here to balance the speed versus accuracy of the D3-MPC computation convergence, we set the stopping criteria of $||x - Ez||_\infty = \ldots$
max_{i \in \mathbb{N}} \left[ \|x_i - E_i z\|_\infty \right] \leq 3.5 \times 10^{-5}. In Fig. 6 we further present the convergence curves of ADMM for different values of $\rho$. Based on these experimental results, we selected $\rho = 100$ for our computation.

The results for voltage performance of D3-MPC scheme for the same prediction profile and prediction errors as in case of C-MPC are depicted in Fig. 7, which clearly establishes that the proposed D3-MPC scheme is also capable of maintaining the desired voltages performance. It is important to note here that in the D3-MPC scheme, there is no centralized optimization performed (control agents at each bus perform their own local optimizations and exchange certain information only among their neighbors), and also instead of assuming the known values of admittances, the SCADA/PMU measurements of system variables $\{V_i, V_k, \theta_i, \theta_k, P_{ik}, Q_{ik}, P_{ki}, Q_{ki} | k \in \mathbb{N}_i\}$ at each bus $i \in \mathbb{N}$ are leveraged to estimate $\{G_{ik}, B_{ik} | k \in \mathbb{N}_i\}$ at that bus. In Figs. 8-9, we present the estimated values of $\{G_{ik}, B_{ik} | i \in \mathbb{N}, k \in \mathbb{N}_i\}$ along with their true values for 8 different lines for representation. These errors can be attributed to the errors in measurement data, but our D3-MPC approach is found to be robust against such modeling errors as demonstrated next.

To explicitly compare the C-MPC versus D3-MPC performance, we computed the deviation in voltage values for each individual bus over time, and present those in Fig. 10. For all 3 cases of prediction error, the voltage deviations between C-MPC and D3-MPC are insignificant in the order of $10^{-3}$ p.u. It can be concluded that at each time step, the ADMM-based proposed D3-MPC scheme reached very close to the global optimum solution, implying that the proposed D3-MPC is robust against the errors introduced due to (a) data-driven estimation of admittances, and (b) the stopping criteria employed for ADMM convergence.

The efficacy of the proposed D3-MPC control strategy is further compared with the conventional droop volt-var control VVC scheme recommended by IEEE 1547. The VVC design is decentralized and solely depends on the local voltage infor-
mation of the particular bus where the controller is located. The voltage profiles for D3-MPC versus VVC schemes are presented in Fig. 11, which clearly indicates the superiority of D3-MPC over VVC.

![Fig. 11. Comparison of voltage profiles under D3-MPC vs. VVC schemes](image)

B. Test Case of IEEE 57 Bus System

For a further validation, we develop results for the IEEE-57 bus system. We modified the standard IEEE 57 bus network, that has generators only at bus-3, 8 and 12, by adding generators also at buses from 13 to 57. Here again, the generations are taken to be a mix of conventional and renewable ones, with renewable generations serving approximately 50% of the total load at its peak availability. The predictions of renewable generation and load profiles are shown in Fig. 12 and taken from real-world data of [28]. Similar to the IEEE 30 bus network, the voltages drop below the desired margin in the absence of any control as depicted in Fig. 13, thus necessitating some reactive control compensation for mitigating the voltage drop.

![Fig. 12. Predicted renewable generation (A), and load profile (B)](image)

The voltages profiles under C-MPC versus D3-MPC control schemes under 5% prediction errors are shown in Fig. 14—both are successful in maintaining the voltages within the margins around the reference. Here to maintain a balance between speed and accuracy, we set stopping criteria to be 

$$\|x - E_z\|_{\infty} = \max_{i \in N} \left[\|x_i - E_i z\|_{\infty}\right] \leq 10^{-4}.$$  

We also present in Fig. 15 the voltage deviations between the respective bus voltages under C-MPC versus D3-MPC schemes. The computed voltage deviations between C-MPC and D3-MPC are in the range of \([-6 \times 10^{-3}, 8 \times 10^{-3}\]), validating that the D3-MPC framework has again achieved close to centrally optimal performance.

![Fig. 13. IEEE 57 Bus: Voltages in the absence of control compensation](image)

![Fig. 14. Comparison of voltage profiles under C-MPC vs. D3-MPC schemes](image)

![Fig. 15. Voltage deviations between the C-MPC and D3-MPC schemes](image)

VI. CONCLUSIONS

In this paper, we developed for the first time a data-driven distributed predictive control method for voltage regulation in wide area networks subject to time-varying generations and consummations. An ADMM-based distributed architecture with the data-driven estimation of system parameters (D3-MPC) is proposed, and its efficacy is validated. The proposed design does not require any a priori knowledge of the line admittances; rather, it is estimated online utilizing the SCADA/PMU measurement data. The test results applied to IEEE 30-bus and 57-bus systems validated the performance and robustness against model/generation/demand uncertainties of the proposed D3-MPC scheme compared to the centralized version C-MPC. We also demonstrated that the proposed distributed D3-MPC offers the level of optimality of the centralized C-MPC while also offering—owing distributed computation—scalability and reduced vulnerability to network
attacks. The proposed D3-MPC is also shown to be superior in performance compared to the standard decentralized practice of VVC (volt-var control), which, while scalable and less prone to attacks due to being decentralized, is quite suboptimal. The proposed D3-MPC thus offers the attractive features of both C-MPC and VVC, and additionally, unlike C-MPC or VVC, it does not require the a priori knowledge of system parameters—rather estimates those in real-time using the SCADA/PMU data. Future research can explore the integration of data-driven learning-based methods into D3-MPC.

References

[1] G. Valverde and T. Van Cutsem, “Model predictive control of voltages in active distribution networks,” *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2152–2161, 2013.

[2] P. N. Vovos, A. E. Kiprakis, A. R. Wallace, and G. P. Harrison, “Centralized and distributed voltage control: Impact on distributed generation penetration,” *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 476–483, 2007.

[3] A. Chakrabortty and P. P. Khargonekar, “Introduction to wide-area control of power systems,” in *2013 American Control Conference*, 2013, pp. 6758–6770.

[4] A. Chakrabortty, *Wide-Area Control of Power Systems*. London: Springer London, 2020, pp. 1–8. [Online]. Available: https://doi.org/10.1007/978-1-4471-5102-9_100125-1

[5] S. Kar, G. Hug, J. Mohammadi, and J. M. Moura, “Distributed state estimation and energy management in smart grids: A consensus + innovation approach,” *IEEE Journal of selected topics in signal processing*, vol. 8, no. 6, pp. 1022–1034, 2014.

[6] J. W. Simpson-Porco, F. Dörfler, and F. Bullo, “Voltage stabilization in microgrids via quadratic droop control,” *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1239–1253, 2017.

[7] ENTSO-E Requirements for grid connection applicable to all generators 8 march 2013. https://epublibdownloads.entsoe.eu/clean-documents/ pre2015/resources/RIG/130308_Final_Version_NC_RIG.pdf.

[8] M. Farivar, X. Zhao, and L. Chen, “Local voltage control in distribution systems: An incremental control algorithm,” in *2015 IEEE international conference on smart grid communications (SmartGridComm)*. IEEE, 2015, pp. 732–737.

[9] X. Zhou and L. Chen, “An incremental local algorithm for better voltage control in distribution networks,” in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 3206–3212.

[10] K. E. Antoniadou-Plytaria, I. N. Kouveliotis-Lysikatos, P. S. Georgilakis, and N. D. Hatziargyriou, “Distributed and decentralized voltage control of smart distribution networks: Models, methods, and future research,” *IEEE Transactions on smart grid*, vol. 8, no. 6, pp. 2999–3008, 2017.

[11] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, “A survey of distributed optimization and control algorithms for electric power systems,” *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

[12] N. Patari, V. Venkataramanan, A. Srivastava, D. K. Molzahn, N. Li, and A. Annaswamy, “Distributed optimization in distribution systems: Use cases, limitations, and research needs,” *IEEE Transactions on Power Systems*, pp. 1–1, 2021.

[13] H. Sun, Q. Guo, J. Qi, V. Ajjarpur, R. Bravo, J. Chow, Z. Li, R. Moghe, E. Nasr-Azadani, U. Tamrakar, G. N. Taranto, R. Tonkoski, G. Valverde, Q. Wu, and G. Yang, “Review of challenges and research opportunities for voltage control in smart grids,” *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 2790–2801, 2019.

[14] O. Lateef, R. G. Harley, and T. G. Habetler, “Bus admittance matrix estimation using phasor measurements,” in *2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2019, pp. 1–5.

[15] S. Bolognani and S. Zampieri, “A distributed control strategy for reactive power compensation in smart microgrids,” *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2818–2833, 2013.

[16] S. Bolognani, R. Carli, G. Cavraro, and S. Zampieri, “Distributed reactive power feedback control for voltage regulation and loss minimization,” *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 966–981, 2014.

[17] P. Šulc, S. Backhaus, and M. Chertkov, “Optimal distributed control of reactive power via the alternating direction method of multipliers,” *IEEE Transactions on Energy Conversion*, vol. 29, no. 4, pp. 968–977, 2014.

[18] B. Robbins and A. D. Domínguez-García, “Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems,” *IEEE Transactions on Power Systems*, vol. 31, pp. 2903–2913, 2016.

[19] B. Zhang, A. Y. Lam, A. D. Domínguez-García, and D. Tse, “An optimal and distributed method for voltage regulation in power distribution systems,” *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1714–1726, 2015.

[20] H. J. Liu, W. Shi, and H. Zhu, “Distributed voltage control in distribution networks: Online and robust implementations,” *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6106–6117, 2017.

[21] E. Dall’Anese, H. Zhu, and G. B. Giannakis, “Distributed optimal power flow for smart microgrids,” *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, 2013.

[22] M. Todescato, J. W. Simpson-Porco, F. Dörfler, R. Carli, and F. Bullo, “Online distributed voltage stress minimization by optimal feedback reactive power control,” *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1467–1478, 2018.

[23] G. Qu and N. Li, “Optimal distributed feedback voltage control under limited reactive power,” *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 315–331, 2020.

[24] C. Conte, T. Summers, M. N. Zeilinger, M. Morari, and C. N. Jones, “Computational aspects of distributed optimization in model predictive control,” in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 6819–6824.

[25] S. Boyd, P. Parikh, and E. Chu, Distributed optimization and statistical learning via the alternating direction method of multipliers. Now Publishers Inc, 2011.

[26] CAISO. California independent system operator. https://www.caiso.com/TodaysOutlook/Pages/supply.html.

[27] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, “A review of deep learning for renewable energy forecasting,” *Energy Conversion and Management*, vol. 198, pp. 11799, 2019.

[28] X. Liu, Z. Zhang, and Z. Song, “A comparative study of the data-driven day-ahead hourly provincial load forecasting methods: From classical data mining to deep learning.” *Renewable and Sustainable Energy Reviews*, vol. 119, p. 109632, 2020.

[29] Y. Guo, Q. Wu, H. Gao, X. Chen, J. Østergaard, and H. Xin, “MPC-based coordinated voltage regulation for distribution networks with distributed generation and energy storage system,” *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 1731–1739, 2018.

[30] S. Samar, S. Boyd, and D. Gorinevsky, “Distributed estimation via dual decomposition,” in *2007 European Control Conference (ECC)*. IEEE, 2007, pp. 1511–1516.

[31] U. DOE, “Ferc,” steps to establish a real-time transmission monitoring framework for emerging grid business structures,” feb. 2006.

[32] Y. Moriz, R. Nishihara, S. Wang, A. Tumanov, R. Liaw, E. Liang, M. Ellob, Z. Yang, W. Paul, M. I. Jordan, et al., “Ray: A distributed framework for emerging [AI] applications,” in *13th USENIX Symposium on Operating Systems Design and Implementation (OSDI 18)*, 2018, pp. 561–577.