Dispatch of Virtual Inertia and Damping: Numerical Method with SDP and ADMM
Tong Han, David J. Hill

Abstract—Power grids are evolving toward 100% renewable energy provided by inverters. Virtual inertia and damping provided by inverters are essential to synchronism and frequency stability of future power grids. This paper numerically addresses the problem of dispatch of virtual inertia and damping (DID) among inverters in the transmission network. The DID problem is formulated as a nonlinear program (NLP) by the Radua numerical method which is flexible to handle various types of disturbances and bounds constraints. Since the NLP of DID is highly non-convex, semi-definite programming (SDP) relaxation for the NLP is further derived to tackle the non-convexity, followed by its sparsity being exploited hierarchically based on chordality of graphs to seek enhancement of computational efficiency. Considering high dimension and inexactness of the SDP relaxation, a feasibility-embodied distributed approach is finally proposed under the framework of alternating direction method of multipliers (ADMM), which achieves parallel computing and solution feasibility regarding the original NLP. Numerical simulations carried out for five test power systems demonstrate the proposed method and necessity of DID.

Index Terms—inverter, virtual inertia, damping, SDP relaxation, sparsity, distributed optimization, ADMM

NOMENCLATURE

A. Scalars

\( V_i \)  
Voltage magnitude of bus \( i \).

\( p_i \)  
Mechanical power input for \( i \in \mathcal{N}_s \) power setpoint for \( i \in \mathcal{N}^s \), negative of load power independent of frequency for \( i \in \mathcal{N}_l \) for \( i \in \mathcal{N}_o \).

\( b_{j_1,j_2} \)  
Susceptance of branch \((j_1,j_2)\).

\( m_i, d_i \)  
Inertia/damping coefficient of generator \( i \).

\( d_{\text{eq}} \)  
Frequency coefficient of load \( i \).

\( b_{\text{eq}} \)  
Equivalent short-circuit susceptance when ignoring the short-circuit resistance.

\( J_i \)  
Component of objective function \( J \) corresponding to disturbance \( k \) and time element \( t \).

\( j \)  
Approximation of \( J \) in (P1).

\( \beta_{\text{cf}} \)  
A proper large number to guarantee positive definiteness of \( A_{\text{d},i} + \beta_{\text{cf}} I \).

\( \hat{\rho} > 0 \)  
The penalty parameter.

\( \sigma_{c_i}^{r \omega}, \sigma_{c_i}^{s \omega} \)  
The 1-th singular value of matrix \( Z_i^{(c_i)} \).

\( r_{\text{abs}}, \epsilon_{\text{rel}} \)  
Primal and dual residuals at iteration \( \kappa \), respectively.

\( \varphi_s, \tilde{\varphi}_s \)  
Slack variables.

\( Z_i^{\text{ml}} \)  
The entry in the 1-th row and 2-th column of \( Z_j^{\text{ml}} \).

This work was supported by the Research Grants Council of the Hong Kong Special Administrative Region through the General Research Fund under Project No. 17209419.

T. Han and D. J. Hill are with the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong (email: han tong@eee.hku.hk, dhill@eee.hku.hk).

B. Vectors

\( \theta \)  
Phase angle of all buses.

\( \omega \)  
Angular frequency of all generators.

\( \omega_{l_i}, \theta_{l_i} \)  
Initial values of \( \omega^k \) and \( \theta^k \).

\( \delta \)  
Lower/upper frequency bound.

\( p_k \)  
Upper bound of angle differences.

\( \omega^k, \theta^k \)  
Lower/upper active power limit of generators.

\( p_{g_i}, p_{d_i} \)  
Lower/upper bound of inertia coefficients.

\( \omega^k, \theta^k \)  
Lower/upper bound of damping coefficients.

\( \omega^k, \theta^k \)  
Constrains equal to \( \theta_{l_i} \) and \( \omega_{l_i} \).

\( [x]^k_i \)  
Sub-vector of \( x \) related to disturbance \( k \) and time element \( i \), i.e., \( [x]^k_i = \text{col}(M^1, D^1, t^k_i, \omega^k_i, \theta^k_i) \).

\( \alpha, \beta, \kappa \)  
\( \text{col}(\alpha^k_{r_1,i}, \beta^k_{r_1,i}, \kappa^k_{r_1,i}) \), with \( k \in D \), \( i \in \mathbb{T}_k \), \( r \in \{0, ..., n_c\} \), \( i \in B \), \( \alpha^k_{r_1,i} \in \mathbb{R}^d \), \( \beta^k_{r_1,i} \in \mathbb{R}^d \), \( \kappa^k_{r_1,i} \in \mathbb{R}^d \).

\( \text{diag}(\varsigma_i) \)  
Arbitrary disjoint sub-vector of \( \text{diag}(\varsigma) \) satisfying \( \text{diag}(\varsigma)^T \text{diag}(\varsigma)^T = \text{diag}(\varsigma)^T \).

\( \text{upper}_i \)  
Analogous to \( \text{diag}(\varsigma) \).

C. Matrices

\( M, D, D_l \)  
\( \text{diag}(m_{i|\in \mathcal{N}_s}) \), \( \text{diag}(d_{i|\in \mathcal{N}_l}) \), \( \text{diag}(d_{i|\in \mathcal{N}_i}) \).

\( B \)  
\( \text{diag}(V_{j_1,j_2} b_{j_1,j_2,i|\in \mathcal{N}} \in \mathbb{B}) \).

\( E_g, E_i \)  
Incidence matrices showing the relationship between \( \mathcal{N}_g \) and \( \mathcal{N} \) and \( \mathcal{N}_i \) and \( \mathcal{N}_i \).

\( E_{a_1}, E_{a_2} \)  
Incidence matrices showing the relationship between \( \mathcal{N}_a \) and \( \mathcal{N}_a \) and \( \mathcal{N} \) and \( \mathcal{B} \).

\( \Omega^k_{11}, \Theta^k_{11} \)  
Collocation coefficient matrices for profile of \( \omega^k \) and \( \theta^k \) at time element \( i \).

\( E_g, G, \hat{\Omega}, \hat{G}_l \)  
\( \text{col}(E_{a_1}, E_{a_2}), \text{diag}(M, \ldots, M), \text{diag}(D, \ldots, D) \).

\( B^k_{\text{ir}}, B^k_{\text{i}r} \)  
\( \text{diag}(\Lambda^k_{\text{ir}}, \Lambda^k_{\text{i}r}) \).

\( \Lambda^k_{\text{ir}}, \Lambda^k_{\text{i}r} \)  
\( \text{diag}(\text{cos}(\Lambda^k_{\text{ir}} \theta^k_0), \sin(\Lambda^k_{\text{ir}} \theta^k_0) - \Lambda^k_{\text{i}r} \theta^k_0) \).

\( \text{equal to } \text{diag}(\varsigma) \) for \( \text{diag}(\varsigma) \) with only the row corresponding to \( \text{diag}(\varsigma) \) remaining and others replaced by 0.

\( [X]^k_i \)  
Principal submatrix of \( X \) related to disturbance \( k \) and time element \( i \), given as \( [x]^k_i [x]^k_i \).

\( Q^k \)  
Sub-matrix of \( Q^k \) by removing the last 2 block rows.
The feasible region defined by inner constraints with \((k, i) \in \Xi_s\) and coupling constraints \((k, i) \in \Xi_s\) satisfying \((k, i - 1) \in \Xi_s\) or \(i = 1\).

The feasible region of \(\Xi_s\).

**I. INTRODUCTION**

**FUTURE** power grids will feature high penetration of renewable energy to mitigate global climate change. Interconnected generation will increasingly substitute for conventional synchronous generators, which however, can cause critical synchronism and frequency stability problems due to low, time-varying and heterogeneous system inertia [1], [2]. Inertia emulation and droop control have been proposed as effective remedies to improve the synchronism and frequency performance of future power grids [3], [4].

Differently to synchronous generators whose inertia and damping are inherent physical properties, virtual inertia and damping of inverters are control parameters of control loops and thus are tunable. Following conventional power grids, inertia and damping coefficients would be set as constants. However, due to high heterogeneity in operating conditions of future power grids causing by wide variations on both generation and demand sides, a fixed setting of virtual inertia and damping is potentially unable to guarantee synchronism and frequency stability under all operating conditions. More importantly, system performances with the fixed setting would probably be far from the optimal with inertia and damping being different for each operating condition.

Naturally inspired by economic dispatch in power grids, we define the concept of dispatch of virtual inertia and damping (DID): short-term determination of the optimal distribution of virtual inertia and damping among a number of inverters, to improve the synchronism and frequency performance, at the lowest control efforts, subject to operational constraints. Compared to economic dispatch with the aim to determine the steady-state optimal active power generation in terms of economic (generation cost), the aim of DID is to determine the optimal active power generation during dynamic processes mainly regarding security (synchronism and frequency stability). It is pointed out that some novel control mechanisms for the fast dynamics have been also developed as an improved alternative to the virtual inertia solution, such as the dynamic-c droop control [5], [6]. Nevertheless, while there is still substantial legacy generation and controls, the DID process will be useful for years to come.

Although the DID problem has not been explicitly proposed previously, researchers have explored some analogues of DID recently [7]–[15]. These studies, known as placement or design of virtual inertia and primary control, mainly focus on optimization of virtual inertia and damping from a planning perspective. Whether from the operating or planning perspective, essentially the problem is the same, that is, an optimal control problem for time-invariant system parameters, i.e., inertia coefficients and damping coefficients. Five aspects are of concern, namely disturbances, system dynamic models, performance metrics, constraints and solution approaches. For
the first aspect, power-step disturbances [8]–[12], [14], power-impulse disturbances [9], [13] and unit-variance stochastic white noise power disturbances [13] were considered to model faults such as generator outages, fluctuations in renewable generation and load steps. For system dynamic models, all existing work applied linear dynamic system models by linearization [8], [9], [12], [14] or utilizing DC power flow [11], [13]. Inverter dynamics were modelled with different fidelity, from the system-level [8], [10], [13], [14] to the device-level [9], [11], [12]. Dependency of loads on frequency was considered in most work [10]–[14]. For measuring system performance, different metrics or their combination were used, including the damping ratio [12], [14], the frequency overshoot [12], [14], the frequency nadir [9], [11], the Rate of Change of Frequency (RoCoF) [9], [11], [12], and $H_2$ norms of linear systems with time-integrated quadratic forms in the voltage angle, frequency deviation, RoCoF or control efforts [8], [9], [11], [13]. Regarding constraints, approximate inverter power limits [9], [11], [12], inertia and damping coefficient bounds [9]–[11], [13], [14] and constraints on some of the performance metrics [12], [14] were considered. The above four aspects determine the problem formulation which is generally a non-convex and large-scale optimization model. Gradient-based optimization methods [9]–[11], [13] and sequential linear programming methods [12], [14] were applied to find a local optimum. Analytical solutions of the optimization model were also derived under restrictive assumptions [8], [13].

Existing studies are incomplete in two aspects. Firstly, the linear power flow model and time-invariant bound constraints were adopted in most previous formulations. Therefore, these formulations are incapable of handling some practical issues, e.g., time-varying frequency bounds and system performance under large disturbances. Secondly, most solution approaches neglected the non-convexity of optimization models, which probably produce inferior optimal solutions regarding global optimality. In light of the above incompleteness, this paper addresses the DID problem from a numerical perspective, aiming to develop a flexible DID formulation and a computationally efficient solution method to obtain near-globally optimal solutions. Specifically, the main contributions of this paper are fourfold: (i) a nonlinear programming (NLP) formulation of DID is developed by the Radua collocation method, where various forms of disturbances and constraints can be easily taken into account; (ii) semi-definite programming (SDP) relaxation for the NLP of DID is derived to address non-convexity of the NLP model; (iii) to improve computational efficiency, sparsity in the SDP relaxation is exploited hierarchically based on the clique decomposition approach and the proposed proposition regarding chordality of graphs; and (iv) under the framework of alternating direction method of multipliers (ADMM), a feasibility-embedded distributed approach is proposed to solve the SDP relaxation for DID in parallel and with solution feasibility to the original NLP being guaranteed.

The rest of this paper contains Section II to Section V organized as shown in Fig. 1 which also gives the purpose of each section, Section VI giving the results of DID of five test systems with our proposed model and solution approach, and Section VII making a conclusion and a prospect for future works.

II. PROBLEM FORMULATION

In this section, system dynamic models and types of disturbances adopted in the DID problem are first introduced. Then, the dynamic optimization model of the DID problem is formulated, which is further transcribed into a tractable finite-dimensional NLP by the Radua collocation method.

A. System Dynamic Model and Disturbances

Consider a transmission network that consists of $n_g$ generation buses including $n_s$ synchronous generators and $n_v$ grid-forming inverters, $n_l$ load buses and $n_o$ buses with neither generator nor load. Denote by $N_g$, $N_v$, $N_{vl}$ and $N_{ov}$ the sets of these five types of buses, respectively. The set of all the buses is denoted by $N$ with $\vert N \vert = n_a$. We further divide $N_v$ into $N_{v_se}, N_{v_dd}, N_{v_dm}$ and $N_{v_de}$, which contain $n_{v_se}$ inverters with fixed damping and inertia, $n_{v_dd}$ inverters with adjustable damping but fixed inertia, $n_{v_dm}$ inverters with adjustable inertia but fixed damping, and $n_{v_de}$ inverters with adjustable damping and inertia, respectively.

Dynamic Model. The structure-preserving model is widely employed for dynamic analysis of transmission networks with only synchronous generators and frequency-dependent loads [16]. By further embedding grid-forming inverter dynamics which are essentially the same as that of synchronous generators in the context of this work [9], and algebraic equations for buses in $N_o$, system dynamics can be modelled by the matrix-form semi-explicit DAEs as follows:

$$\begin{cases}
E_g \dot{\theta} = \omega \\
\dot{\omega} = -M^{-1}D\omega + M^{-1}E_g\tilde{p} - M^{-1}E_gE_aB\sin(E_n^T\theta) \\
E_l \dot{\theta} = D_l^{-1}E_l\tilde{p} - D_l^{-1}E_lE_nB\sin(E_n^T\theta) \\
0 = E_o\tilde{p} - E_oE_nB\sin(E_n^T\theta)
\end{cases}
$$

where $\tilde{p} = \text{col}(p_i - V_i^2b_i) \in \mathbb{R}^{n_a}$ with $i \in N$ and $V_i$ assumed to be constant in the study of angle stability and frequency stability.

Disturbances. Four typical types of disturbances, including the power-step disturbance, power-ramp disturbance, power-fluctuation disturbance and three-phase short circuit are considered for DID. The disturbance set is denoted by $D = D_1 \cup D_2 \cup D_3 \cup D_4$, with $D_1$ to $D_4$ representing the sets of each type of disturbances, respectively.
B. DO Formulation of DID

Considering the disturbances $\mathcal{D}$ and system dynamics over finite-time horizon $t \in [t_0, t_f]$, the DID problem can be formulated as the dynamic optimization (DO) model as follows:

\[
\begin{align*}
\text{(P1)} & \quad \min_{M,D,\theta^k,\omega^k} J = \sum_{k \in \mathcal{D}} J_k \\
& \quad \text{s. t. } \forall k \in \mathcal{D} \\
& \quad \omega^k(t_0) = \omega_{t_0} \\
& \quad E_0\theta^k(t_0) = E_0\theta_{t_0} \\
& \quad \omega^k \leq \omega^k \leq \omega^k \\
& \quad -\delta \leq E_n^t\theta^k \leq \delta \\
& \quad p_g \leq E_0p^k - M\omega^k - D\omega^k \leq p_g \\
& \quad m \leq M1 \leq m \\
& \quad d \leq DL \leq d
\end{align*}
\]

with

\[
J_k = \int_{t_0}^{t_f} \left( L_p(\theta^k,\omega^k) + L_c(M,D,\theta^k,\omega^k) \right) dt
\]

\[
L_p(\theta^k,\omega^k) = \theta^k W_1^k + \omega^k W_2^k + \omega^k W_3^k + \theta^k W_4^k
\]

\[
L_c(M,D,\theta^k,\omega^k) = (M\omega^k + D\omega^k)W_1^k(W_2^k + D\omega^k) + \theta^k W_3^k W_5^k
\]

where $J$ is an integral cost function consisting of system performance term $L_p$ and control effort term $L_c$ for each disturbance; $L_p$ and $L_c$ are both in squared $H_2$-norm forms, in which $L_p$ measures phase cohesiveness, frequency boundedness and frequency oscillation, and $L_c$ measures control efforts of generators and loads; (2a) are system DAE constraints; (2c) and (2d) are initial value constraints for differential variables; (2e) are system frequency constraints; (2f) are transient disturbance; (2g) are finite-time horizon constraints; (2h) are state, input and output constraints.

Remark 1. Since this work focuses more on solution approach for the DID model, we adopt simple system dynamic models and ignore some practical issues in the DID model. For example, the governor dynamics of synchronous generators are ignored, and limits on the maximal value of the rate of change of frequency and total energy released by inverters in response to a disturbance are not considered. However, the proposed approach in the following is still applicable to the case where these practical issues are considered since they generally cause new non-convexity and nonlinearity.

C. NLP Formulation of DID

The DO formulation of DID, being intractable to be solved directly, is further transcribed into a finite-dimensional NLP by collocation methods. Seeing that slow dynamics of synchronous generators and fast dynamics of inverters with small inertia coefficients can constitute stiff DAEs, Radua collocation is employed in our work [18]. This method corresponds to the fully implicit Runge-Kutta method and has similar high-order accuracy and stability properties. For stiff power grids, Radua method is L-stable. Furthermore, in Radua collocation, endpoints are collection points, which allows constraints to be set easily at the end of each element [18].

Specifically, for each disturbance $k \in \mathcal{D}$, the time interval $[t_0, t_f]$ is first divided into $n_k^i$ finite elements of length $h_k^i$, denoted as $\mathcal{T}_k^i = \{1, ..., n_k^i\}$, such that $\sum_{i=1}^{n_k^i} h_k^i = t_f - t_0$.

The solution of DAEs can be approximated by Lagrange polynomials over each element. For any given $k \in \mathcal{D}$ and $i \in \mathcal{T}_k^i$, we have

\[
t = t_{i-1} + h_k^i \tau
\]

\[
\hat{\omega}_k(t) = \Omega_k^i \ell(\tau) = \ell(\omega(\tau))\omega_k^i
\]

\[
\hat{\theta}_k(t) = \Theta_k^i \ell(\tau) = \ell(\theta(\tau))\theta_k^i
\]

where $\omega_k(t) \in \mathbb{R}^{n_k}$ and $\theta_k(t) \in \mathbb{R}^{n_k}$ are vectors of piecewise polynomial with $(n_k+1)$ degree, approximating $\omega_k(t)$ and $\theta_k(t)$, respectively; $\Omega_k^i = \text{row}(\omega(\tau))_{i} \in \mathbb{R}^{n_k \times (n_k+1)}$ and $\Theta_k^i = \text{row}(\theta(\tau))_{i} \in \mathbb{R}^{n_k \times (n_k+1)}$ with $\tau \in [0, ..., n_k]$, are vectorization of collocation coefficient matrices; $\ell(\tau) = \ell(\tau)_{i} \in \mathbb{R}^{n_k}$ with $\tau \in [0, ..., n_k]$ and $\ell(\tau)$ being the Lagrange polynomial with order $(n_k+1)$.

\[
\ell(\tau) = \sum_{r=0}^{n_k} \tau - \tau_r \frac{\tau - \tau_r}{\tau_j - \tau_r}
\]

with $\tau_r$ representing location of collocation points within each time element, $t_0 = 0$, $\tau_{n_k} = 1$, and $\tau_j < \tau_{j+1}$, $j \in [0, ..., n_k - 1]$.

$\ell(\tau) = \ell(\tau)_{i} \otimes I_{n_k}$ and $\ell(\tau) = \ell(\tau)_{i} \otimes I_{n_k}$. For the convenience of NLP formulation, we mainly used the second expression of $\hat{\omega}_k$ and $\hat{\theta}_k$. Additionally, we will use the derivative of $\ell(\tau)$, and in matrix form, we have $\hat{\ell}(\tau) = \hat{\ell}(\tau)_{i} \otimes I_{n_k}$ and $\ell(\tau) = \hat{\ell}(\tau)_{i} \otimes I_{n_k}$; with $\hat{\ell}(\tau) = \frac{d\ell(\tau)}{dt}$, $\frac{d\ell(\tau)}{dt} = \frac{d\ell(\tau)}{dt}$.

With $\omega_k$ and $\theta_k$ replaced by $\hat{\omega}_k$ and $\hat{\theta}_k$ respectively, $J_k$ can be approximated by $J_k^i$ given as

\[
J_k^i = \int_{t_{i-1}}^{t_i + h_k^i} L_p(\hat{\theta}_k,\hat{\omega}_k) + L_c(M,D,\hat{\theta}_k,\hat{\omega}_k) dt
\]

\[
= h_k^i \int_{0}^{1} L_p(\ell(\tau)_{i} \ell(\tau))\hat{\omega}_k + L_c(M,D,\ell(\tau)_{i} \ell(\tau))\hat{\omega}_k dt
\]

Since $L_p$ and $L_c$ are both in the form of polynomial w.r.t time $t$ or $\tau$, definite integral of them can be computed analytically and thus, $J_k^i$ can be explicitly formulated as a polynomial function in terms of $M$, $D$, $\theta^i_k$ and $\omega^i_k$, given by

\[
J_k^i(M,D,\theta^i_k,\omega^i_k) = \theta^i_k W_1^i + \theta^i_k W_2^i + \theta^i_k W_3^i + \theta^i_k W_4^i + \theta^i_k W_5^i + \omega^i_k W_1^i + \omega^i_k W_2^i + \omega^i_k W_3^i + \omega^i_k W_4^i + \omega^i_k W_5^i
\]
where matrices $W^k_{ij}$ to $W^k_{ii}$ are given in Table I.

 Lagrange polynomial $\ell_j$ satisfies $\ell_j(\tau_r) = \delta_{jr}$ for $\forall r, j \in \{0, \ldots, n_c\}$, where $n_c$ is the number of collocation points. With this property, substituting the polynomial into DAE constraints (2b) and enforcing the resulting algebraic equations at the interpolation points $\tau_r$, lead to the collocation equations for DAE constraints as follows:

$$
\begin{align*}
E_i s \ell_j(\tau_r) \theta^k_i &= h^k_i \omega^k_i, \quad \forall r \in \{1, \ldots, n_c\} \\
\ell_j(\omega^k_i) &= h^k_i M^{-1} D\omega^k_i + h^k_i M^{-1} E_i g^k \tau_i \\
E_i \ell_j(\omega^k_i) \dot{\theta}^k_i &= h^k_i D^{-1} E_i \tau_i \\
- h^k_i \theta^k_i E_i n + B^k_i \sin(E_i T \theta^k_i) \quad \forall r \in \{1, \ldots, n_c\} \\
0 &= -E_i \theta^k_i - E_i B^k_i \sin(E_i T \theta^k_i) \quad \forall r \in \{0, \ldots, n_c\}
\end{align*}
$$

(10)

Additionally, $\omega^k_0$ and $E_i \theta^k_0$ are determined by initial conditions (2c) and (2d), or enforced by the continuity of the differential variable profiles across element borders as follows:

$$
\begin{align*}
\omega^k_0 &= \omega^k_{t_0} \\
E_i \theta^k_0 &= \theta^k_{t_0}
\end{align*}
$$

(11)

(12)

Enforcing path constraints (2e) to (2g) at all collocation points gives

$$
\begin{align*}
\omega^k_{t_r} \leq \omega^k_{t_r} &\leq \omega^k_{t_r} \quad \forall r \in \{1, \ldots, n_c\} \\
-\delta \leq \theta^k_{t_r} &\leq \delta \quad \forall r \in \{1, \ldots, n_c\} \\
p_r \leq E_i p^k_{r} \quad \forall r \in \{1, \ldots, n_c\}
\end{align*}
$$

(13)

(14)

Finally, by combining (2h), (2i) and (9) to (14), we can derive the NLP formulation of DID given by

$$
\begin{align*}
\text{P2:} \min_{M, D, \omega, \theta} \hat{J}(M, D, \theta, \omega) = \sum_{k \in D, i \in T^k} \hat{J}^k_i(M, D, \theta^k_i, \omega^k_i)
\end{align*}
$$

(15a)

(15b)

(15c)

(15d)

(15e)

(15f)

(15g)

where matrices $C^k_{1i}, C^k_{2i}, A^k_{1i}, A^k_{2i}, H^k_{1i}, H^k_{2i}, L^k_1, L^k_2, b^k_{1i}, b^k_{2i}, c^k_{1i}, c^k_{2i}, c^k_{3i}, c^k_{4i}, c^k_{5i}$ and $c^k_{6i}$ are given in Table I.

### III. SDP RELAXATION FOR THE NLP OF DID

The NLP of DID is non-convex due to not only the quadratic interval constraints and equality constraints that contain sine and quadratic terms, but also the non-convex polynomial objective function. Solving the NLP directly by local optimization approach will very likely produce inferior solutions in terms of global optimality. Therefore, SDP-based convex relaxation of the NLP of DID is derived in this section, aiming to enhance global optimality of final solutions.

#### A. SDP Relaxation of DID with Linearized Power Flow

We first consider a simpler case where power flow terms in (1) are linearized. This case is valid when $D$ contains no major disturbances. For the cost function, we define the lifting as follows:

$$
[l^k_{zi}, l^k_{di}] = [M \omega^k_i, \tilde{D} \omega^k_i] \quad \forall k \in D, i \in T^k.
$$

(16)

which replaces products with non-linear variables. Then problem (P2) with linearization and lifting (16) can be equivalently formulated as the following QCQP problem:

$$
\begin{align*}
\text{P3:} \min_{x} \sum_{k \in D, i \in T^k} [x^k_{zi}]^T P^k_{0i} [x^k_{zi}] \\
\text{s.t.} \col([x^k_{zi}]^T P^k_{0i}, [x^k_{zi}])^T Q^k_{1i} [x^k_{zi}] + b^k_{1i} = 0 \\
A^k_{1i} [x^k_{zi}] \leq b^k_{2i} \\
L^k_{1i} [x^k_{zi}] - L^k_{14} [x^k_{zi}] = 0
\end{align*}
$$

(17a)

(17b)

(17c)

(17d)
Here for \((h, r, j), h \in \{1, 2, 3\}, r \in \{1, \ldots, n_c\}\) for \(h = 1\) and \(\{0, \ldots, n_c\}\) for \(h = 2\) or 3, and \(j \in \mathcal{N}_g\); (17b) corresponds to lifting equalities (16) and (15b) in (P2), (17c) corresponds to (15c) to (15e) in (P2), and (17d) corresponds to (15f) and (15g) in (P2). Note that (15e) in (P2) is in quadratic form but transformed into linear form in (P3). Coefficient matrices in (P3) are given in Table II. By further introducing a matrix variable \(X = xx^T\), (P3) is reformulated as

\[
(P4) \min_{E} \sum_{k \in D_i \in \mathcal{T}^k} \operatorname{Tr}(P_{0i}^k[X]^k) \quad (18a)
\]

s.t. \(\forall k \in D_i \in \mathcal{T}^k\)

\[
\begin{align*}
&\text{col} (\operatorname{Tr}(P_{(h,r,j)}^k|x|^k)) + Q_i^k[x]^k + \bar{b}_i^k = 0 \quad (18b) \\
&A_i^k[x]^k \preceq b_i^k, L_{13}[x]^k - L_{14}[x]^k = 0 \quad (18c) \\
&Z = [X, x^T] \succeq 0 \quad (18d) \\
&\text{rank}(Z) = 1 \quad (18e)
\end{align*}
\]

Dropping the non-convex rank constraint (18e) gives the convex SDP as follows:

\[
(P5) \quad (18a) \sim (18d) \quad (19)
\]

We call (P5) a SDP relaxation of (P3) since the feasible region of (P3) or (P4) is a subset of the feasible region of (P5). If an optimal point \(Z^*\) of (P5) satisfies \(\text{rank}(Z^*) = 1\), i.e., \(X^* = x^*x^*^T\), then \(x^*\) is also the global optimum of (P3).

### Table II

| \(P_{0i}^k\) | Matrices in the QCQP formulation of DID |
|-------------|----------------------------------------|
| \(b_i^k\)   | \[O O O O O O O O O O O O O O \] \[b_i^k\] |
| \(Q_i^k\)   | \[O O C_i^k A_i^k O O O O\] \[b_i^k + C_i^k A_i^k\] |
| \(A_i^k\)   | \[O O C_i^k A_i^k O O O O\] \[b_i^k + C_i^k A_i^k\] |
| \(L_{13}^k\) | \[O O C_i^k A_i^k O O O O\] \[b_i^k + C_i^k A_i^k\] |

Then with the assumption that \(\vec{s} \in \{\frac{\pi}{2}, \pi\}\), (P2) with lifting (16), linearization for \(k \in D_1 \cup D_2 \cup D_3\) and Approximation 1 for \(k \in D_4\) can also be formulated as a QCQP that we call (P6) hereafter. Variables corresponding to \(\alpha, \beta, \zeta\) in Approximation 1, i.e., \(\alpha, \beta, \zeta\), are introduced for (P6). Clearly, \(\alpha\) only appears in linear form, and thus by introducing \(Y = \text{col}(\beta, \zeta) \cdot \text{col}(\beta, \zeta)^T\), we obtain the formulation equivalent to (P6) as follows:

\[
(P7) \min_{Z, \vec{Z}} \sum_{k \in D_i \in \mathcal{T}^k} \operatorname{Tr}(P_{0i}^k[X]^k) \quad (20a)
\]

s.t. \((18b), (18c)\) \(\forall k \in D_1 \cup D_2 \cup D_3, i \in \mathcal{T}^k\) \(\quad (20b)\)

\[
(21a) \sim (21b) \quad \forall k \in D_4, i \in \mathcal{T}^k \quad (20c)
\]

\[
Z = [Y, \beta, \zeta] \succeq 0, Z = 0 \quad (20d)
\]

\[\text{rank}(Z) = 1, \text{rank}(\overline{Z}) = 1 \quad (20e)\]

with \((21a) \sim (21b)\) written as follows:

\[
\begin{align*}
&\operatorname{col}\{\operatorname{Tr}(P_{(h,r,j)}^k[x]^k)\} + C_i^k \operatorname{col}\{\operatorname{Tr}(P_{(h,r,j)}^k[Y])\} = b_i^k = 0 \quad (21a) \\
&\operatorname{col}\{\operatorname{Tr}(P_{(h,r,j)}^k[x]^k)\} = 0 \quad (21b) \\
&\epsilon_i^k \leq \text{col}\{\operatorname{Tr}(P_{(h,r,j)}^k[x]^k)\} \leq \epsilon_i^k \quad (21c) \\
&\bar{A}_i^k[x]^k \preceq \bar{b}_i^k, L_{13}[x]^k - L_{14}[x]^k = 0 \quad (21d) \\
&\epsilon_i \text{col}(\text{Tr}(P_{(h,r,j)}^k[x]^k) + A_i^k \theta^k - \gamma_i) \leq 0, \epsilon_i \bar{A}_i^k \text{col}(\text{Tr}(P_{(h,r,j)}^k[x]^k) + A_i^k \theta^k - \gamma_i) \leq 0 \quad (21e) \\
&\epsilon_i \text{col}\{\operatorname{Tr}(P_{(h,r,j)}^k[x]^k)\} = 0, \epsilon_i \text{col}(\text{Tr}(P_{(h,r,j)}^k[x]^k) + A_i^k \theta^k - \gamma_i) = 0 \quad (21f) \\
&\epsilon_i \text{col}(\text{Tr}(P_{(h,r,j)}^k[x]^k) + A_i^k \theta^k - \gamma_i) = 0, \epsilon_i \text{col}(\text{Tr}(P_{(h,r,j)}^k[x]^k) + A_i^k \theta^k - \gamma_i) = 0 \quad (21g) \\
\end{align*}
\]

B. SDP Relaxation of DID with Nonlinear Power Flow

When disturbance set \(D\) contains large disturbances, linearization of power flow equations can lead to unacceptable approximation errors. Hence fidelity of nonlinearity for disturbances in \(D_3\) should be reserved. We first propose the following quadratic approximation of sine function in domain \([-\theta_c, \theta_c]\) with \(\theta_c \in [\frac{-\pi}{2}, \pi]\). Numerical analysis for this approximation can be found in Appendix-A.

**Approximation 1.** \(\sin \theta \leftrightarrow \beta^T \zeta\) with

\[
\begin{align*}
\begin{bmatrix}
\theta^2 + \sin \theta \theta & \frac{\pi}{2} \theta - 1 \\
\frac{\sin \theta}{\theta} & \theta^2 + \sin \theta \theta - \frac{\pi}{2} \theta + 1
\end{bmatrix} \zeta = 0 \quad \text{with} \quad \phi = \frac{1 - \sin \theta}{\theta} = \left(\frac{\pi}{2} \right)^2 \\
\theta - \theta^2 \zeta = 0, \beta = 1, \zeta = 1, (\beta - 1)^2 = 0, \alpha = E \beta \leq 0, \alpha \geq 0
\end{align*}
\]

where \(\theta \in [-\theta_c, \theta_c]\) with \(\theta_c \in [\frac{-\pi}{2}, \pi]\), \(\beta \in \mathbb{R}^3\), \(\zeta \in \mathbb{R}^3\), \(\alpha \in \mathbb{R}^4\), \(\theta_c \in [0, \frac{\pi}{2}]\) and \(\theta_c\) are known parameters; \(\theta_{\alpha} = [-\theta_c, -\theta_c, \theta_c, \theta_c]^T\); and \(E = [1, 0, 0, [1, 1, 0, [0, 1, 1], [0, 0, 1]]\).

Since (P5) can be regarded as a special case of (P8), only (P8) is considered thereafter.

\[
(P8) \quad (20a) - (20d) \quad (22)
\]
IV. EXPLOITING SPARSITY IN SDP RELAXATION

Common approaches for solving SDPs can only handle multiple small PSD matrices efficiently [19], [20]. In SDP relaxation for DID, the large size of $\mathbf{Z}$ and $\tilde{\mathbf{Z}}$ significantly affects computational efficiency of solution approaches for SDPs. Hence this section further exploits sparsity in (P8) to decompose large-dimensional PSD matrix constraints into smaller ones which can be handled much more efficiently.

A. Clique Decomposition Approach

We will use some basic concepts in graph theory, including maximal clique, chordality, chordal extension and clique tree. Reader should refer to [21] for more details of these concepts. The symmetric matrix $\mathbf{Z}$ can be associated with a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ with $\mathcal{V}$ being the row (or column) index set of $\mathbf{Z}$ and $\mathcal{E} = \{(i, j) | i \neq j, [\mathbf{Z}]_{ij} \text{ features in the objective function or constraints}\}$. Here $\mathcal{E}$ is called the aggregate sparsity pattern of SDP in terms of $\mathbf{Z}$. A matrix $\mathbf{Z}' \in \mathbb{S}^{\mathcal{V}}$ is called a symmetric partially specified matrix. A matrix $\mathbf{Z} \in \mathbb{S}^{\mathcal{V}}$ is called a positive semi-definite completion of $\mathbf{Z}'$ if $[\mathbf{Z}]_{ij} = [\mathbf{Z}']_{ij}$ for all $(i, j) \in \mathcal{E}$ and $\mathbf{Z} \succeq 0$.

Then the clique decomposition approach to decompose $\mathbf{Z}$ can be summarized as the following steps: (i) compute a chordal extension $\mathcal{F} \supseteq \mathcal{E}$ if $\mathcal{G}$ is not chordal; (ii) identity the set of maximal cliques $\mathcal{K} = \{C_1, ..., C_{n_{mc}}\}$ of $\mathcal{G}(\mathcal{V}, \mathcal{E})$ or $\mathcal{G}(\mathcal{V}, \mathcal{F})$ if $\mathcal{G}$ is not chordal; (iii) compute the clique tree $\mathcal{T}(\mathcal{K}, \mathcal{L})$; and (iv) decompose $\mathbf{Z}$ into $\mathbf{S}_c(\mathbf{Z}) \succeq 0, \forall i \in \{1, 2, ..., n_{mc}\}$ with equality constraints introduced to equate the overlapping entries in maximal cliques [19]. Considering the complexity of $\mathbf{Z}$ and $\tilde{\mathbf{Z}}$, we propose Proposition 1 (see Appendix-B) so as to explore sparsity of SDP relaxation in a hierarchical way.

Remark 2. Taking $\mathbf{Z}$ for example, it can be associated with graphs $\mathcal{G}(\mathcal{V}, \mathcal{E})$ and $\tilde{\mathcal{G}}(\tilde{\mathcal{V}}, \tilde{\mathcal{E}})$, at the element and block levels, respectively. Nodes in $\mathcal{V}$ and $\tilde{\mathcal{V}}$ correspond to elements and blocks of $\mathbf{Z}$, respectively. Thus by Proposition 1, steps (i) and (ii) of the clique decomposition approach can be conducted hierarchically.

B. Aggregate Sparsity Pattern and Decomposition of SDP

The results of chordality and maximal cliques of the associated graph of $\mathbf{Z}$ is given by Proposition 2 (see Appendix-B). Denote by $T(\mathcal{K}, \mathcal{L})$ the clique tree for $\mathcal{G}$. Then $T$ only needs to satisfy the two properties as follows: (i) $\forall j \in \mathcal{N}_g$, the induced graph by nodes $\{C_{ej} | C_{ej} \in \bigcup_{k \in D, i \in \mathcal{P}_k} \mathcal{K}_{ei}^{k}\}$ forms a subtree of $T$; and (ii) $\forall k \in D, i \in \mathcal{P}_k$, the induced graph by the node set $\mathcal{K}_{ei}^{k}$ with substitution $C_{ej} \rightarrow C_{nj}$ forms a clique tree of $\mathcal{G}_n$. Other maximal cliques can be arranged arbitrarily provided that $T$ is tree-structured. Equality constraints only need to be defined for overlapping entries between maximal cliques within the induced graphs in properties (i) and (ii).

Remark 3. To decompose $\mathbf{Z} \succeq 0$, we mainly need to compute a chordal extension of $\mathcal{G}_n$, $\mathcal{K}_n$, and a clique tree of the extended $\mathcal{G}_n$. In our work, a chordal extension of $\mathcal{G}_n$ is obtained using a fill-reducing Cholesky factorization of matrix $A_{adj} + \beta_{ef} I$.

The Bron-Kerbosch algorithm [22] is used to identify $\mathcal{K}_n$. The clique tree is obtained from a maximum-weight spanning tree of a graph with nodes corresponding to $C_{nj}$ and edge weights between each node pair given by the number of shared bases in each clique pair. The maximum-weight spanning tree can be computed by the Kruskal’s algorithm [23]. Readers should refer to [21]–[24] for more details of these computations.

Regarding the aggregate sparsity pattern of $\tilde{\mathbf{Z}}$, it is trivial to conclude that $\tilde{\mathbf{Z}}$ is chordal with the set of maximal cliques given by

$$\tilde{\mathcal{K}} = \bigcup_{k \in \mathcal{D}, i \in \mathcal{T}_k^{k}} \tilde{\mathcal{K}}^{ki}_i$$

$$\tilde{\mathcal{K}}^{ki}_{i} = \{C_{ej} | C_{ej} = \tilde{I}^{ki}_j(\beta, r, t, j) \cup \tilde{I}^{ki}_r(\beta, r, t, j) \cup \tilde{I}^{-1}(1), \forall r, t, j \in \tilde{\mathcal{P}}\}$$

where $\tilde{I}^{ki}_j(\beta, r, t, j)$ represents the index of $\tilde{\mathbf{Z}}$ corresponding to the $j$th entries in $\beta^{(r, t)}_i$, with $(r, t) \in \tilde{\mathcal{P}} = \{0, ..., n_c\} \times B \times \{1, 2, 3\}$, and $\tilde{I}^{ki}_r(\beta, r, t, j)$ the index of $\tilde{\mathbf{Z}}$ corresponding to the $j$th entries in $\beta^{(r, t)}_i$, with $(r, t) \in \tilde{\mathcal{P}} = \{0, ..., n_c\} \times B \times \{1, 2, 3\}$. No equality constraints need to be defined since overlapping entries are all $\tilde{S}_{\tilde{I}^{-1}(1)}(\tilde{\mathbf{Z}}) = 1$.

C. Decomposition of the SDP Relaxation

We further introduce symmetric matrix variables $\tilde{\mathbf{Z}}_{c_{ij}} \in \mathbb{R}^{\{c_{ij}\} \times \{c_{ij}\}}$, $\forall C_{ij} \in \mathcal{K} \cup \mathcal{K}_n$. According to step (iv) of the clique decomposition approach, (P8) can be decomposed as

$$\min \left\{ \sum_{C_{ij} \in \mathcal{K}} \sum_{k \in \mathcal{D}, i \in \mathcal{T}_k^{k}} \text{Tr}(P_{c_{ij}} \tilde{\mathbf{Z}}_{c_{ij}}) \right\}$$

subject to $\forall C_{ij} \in \mathcal{K} \cup \mathcal{K}_n$

$$\sum_{c_{ij} \in \mathcal{K} \cup \mathcal{K}_n} \text{Tr}(P_{c_{ij}} \tilde{\mathbf{Z}}_{c_{ij}}) \leq 0$$

for $\gamma = 1, 2, ..., n_g$.

$$\mathcal{S}_{ij}(\tilde{\mathbf{Z}}_{c_{ij}}) = \mathcal{S}_{ij}(\tilde{\mathbf{Z}}_{c_{ij}}) = 0 \quad \forall C_{ij} \in \mathcal{K} \cup \mathcal{K}_n$$

$$\mathcal{S}_{ij}(\tilde{\mathbf{Z}}_{c_{ij}}) = \mathcal{S}_{ij}(\tilde{\mathbf{Z}}_{c_{ij}}) = 0 \quad \forall C_{ij} \in \mathcal{K} \cup \mathcal{K}_n$$

Here in (24d), all intersections between cliques for any given $j \in \mathcal{N}_g$ correspond to the same principal submatrix of $\mathbf{Z}$ defined by index set $\{I(M, k, j), \mathcal{I}(D, l), \mathcal{I}(1)\}$; in (24e), (24f) and (24g) are equivalent to (18b), (18c), (21a) to (21d) and (21e) to (21h), respectively; (24c) and (24d) equate the overlapping entries; (24e) and (24f) are equivalent to the second equations in (18c) and (21d); (24g) is equivalent to (20d). Thereby, large size PSD matrices in (P8), i.e., $\mathbf{Z}$ and $\tilde{\mathbf{Z}}$, are decomposed into multiple much smaller ones, i.e., $\mathbf{Z}_{c_{ij}}$ in (24g), and (P9) can be solved more efficiently than (P8).

V. ADMM-BASED DISTRIBUTED OPTIMIZATION FOR DID

The collocation method endows the DID formulation flexibility to handle various forms of disturbances and bound constraints, while an inevitable consequence is high dimension of (P2) and induced problems. Moreover, the SDP relaxation could be inexact. In this section, we propose a feasibility-embedded distributed approach for solving (P9) to address
these two issues simultaneously under the framework of ADMM. For the high dimension, ADMM is utilized to separate the SDP with intractable size into a series of small-size SDPs which can be solved in parallel. To cope with inexactness of the SDP relaxation, we further embed solution feasibility into the solution process with the idea of the ADMM-based restricted low-rank approximation approach [25]. In this way, near-locally optimal solutions of (P2) are promisingly obtained with moderate computational burdens.

Notaion: To simplify expressions, a special kind of variables and corresponding operations are first introduced. We call \(X=\begin{bmatrix} X_1, X_2, \ldots, X_n \end{bmatrix}^T\) where \(X_i\) are all matrices, a matrix vector. If another matrix vector \(Y=\begin{bmatrix} Y_1, Y_2, \ldots, Y_n \end{bmatrix}^T\) satisfies that \(\forall i=1, 2, \ldots, n, X_i\) and \(Y_i\) are with the same size, then \(X\) and \(Y\) are with the same size; \(X+Y=\begin{bmatrix} X_1+Y_1, \ldots, X_n+Y_n \end{bmatrix}^T\); \(X^T=\begin{bmatrix} X_1^T, \ldots, X_n^T \end{bmatrix}^T\); \(\text{Tr}(X)=\sum_{i=1}^n \text{Tr}(X_i)\) and \(X^{TT}=\begin{bmatrix} X_1^T X_1^T, \ldots, X_n^T X_n^T \end{bmatrix}\). Frobenius norm of \(X\) is defined as \(|X|_F^2=\sum_{i=1}^n \text{Tr}(X_i^T X_i)\). A linear matrix vector function is defined as \(\hat{X}=\begin{bmatrix} \hat{X}_1, \hat{X}_2, \ldots, \hat{X}_n \end{bmatrix}^T\) → \(Z=\begin{bmatrix} Z_1, Z_2, \ldots, Z_n \end{bmatrix}^T\), where \(X\) and \(Z\) are matrix vectors, and \(\forall j=1, 2, \ldots, m, \exists i=1, 2, \ldots, n\) and matrix \(P_j\), let \(Z_i=P_j X_j^T, \text{rank}(X)=\max\{\text{rank}(X_i), i=1, 2, \ldots, n\}\); vec(X) = [vec(X_1)^T, vec(X_2)^T, \ldots, vec(X_n)^T]^T with vec(X_i) being the vectorization of matrix X_i; \(\cal{X}\) is called a symmetric matrix vector or \(\cal{X}\) is symmetric if \(\forall i=1, 2, \ldots, n, X_i\) is a symmetric matrix; diag(X)=[diag(X_1)^T, diag(X_2)^T, \ldots, diag(X_n)^T]^T; lower(X)=[lower(X_1)^T, lower(X_2)^T, \ldots, lower(X_n)^T]^T with lower(X_i) being the vectorization of all the entries under the main diagonal in matrix X_i following a column-major order; and upper(X) is analogous to lower(X) but for all the entries above the main diagonal and following a row-major order.

\[ A. \text{ Distributed Solution Approach} \]

We first present the distributed approach for solving (P9) without considering feasibility of solutions. Separation is conducted by decoupling portions corresponding to different sets of disturbances and time elements. Without loss of generality, it is assumed that all pairs of \((k, i)\) are divided into \(\cal{N}_S\) disjoint sets, i.e., \(\Xi_N=\{\ldots, (k, i), \ldots\}\) with \(s \in \mathbb{P}\), and \(\forall s \in \mathbb{P}, \Xi_N(s)\) define the following sets:

\[ \Xi_{\cal{M}}=\{(k, i)\mid \forall s \in \mathbb{P}\} \]
\[ \Xi_{\cal{M}}=\{(k, i)\mid \forall s \in \mathbb{P}\} \]
\[ M_{\Xi_{\cal{M}}} \equiv \Xi_{\Xi_N(s)} \]

To make (P9) separable regarding variables corresponding to each \(\Xi_N\), the following auxiliary matrix variables are created:

\[ Z_{\Xi_N}^{md} \in \mathbb{R}^3 \quad \forall j \in \mathbb{N}_g \]
\[ Z_{\Xi_N}^{cd} \in \mathbb{R}^{C_{\Xi_N}} \quad \forall (k, i) \in \mathbb{M}_N, C_{\Xi_N} \in \mathbb{K}_{\Xi_N} \]

Then we divide constraints in (P9) into two groups, called inner constraints that involve only variables corresponding to cliques from the same \(\mathbb{K}_{\Xi_N}\), and coupling constraints that involve variables corresponding to cliques from different \(\mathbb{K}_{\Xi_N}\), respectively. Specifically, constraints (24b), (24c) and (24g) are inner constraints and the others are coupling constraints. Furthermore, (P9) can be rewritten as the following form:

\[ \min_{\tilde{Z}\in\Xi_N, \hat{Z}\in\Xi_{\mathbb{P}}} \sum_{s \in \mathbb{P}} \sum_{(k, i) \in \Xi_N} \sum_{\xi \in \xi_c} \sum_{\Xi_N} \text{Tr}(P_c \hat{Z}_{\Xi_N}) \]

s.t. \(\forall s \in \Xi_{\mathbb{P}}\)

where \(C^A\) and \(C^B\) are proper linear matrix vector functions to make constraint (27b) equivalent to the following ones:

\[ S_{\Xi_N}(\hat{Z}_{\Xi_N})-Z_{\Xi_N}^md=0 \quad \forall j \in \mathbb{N}_g, C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k, (k, i) \in \mathbb{M}_N \]

Clearly, (P9) can be regarded as a consensus problem with both global and local variables. For any given \(j \in \mathbb{N}_g\), all \(S_{\Xi_N}(\hat{Z}_{\Xi_N})\) with \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\) and \(s \in \mathbb{P}\), achieve consensus globally; for any given \(k, i \in \mathbb{M}_N\), \(S_{\Xi_N}(\hat{Z}_{\Xi_N})\) with \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\) and \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\), achieve consensus locally; and for any given \(k, i \in \mathbb{M}_N\) and \(j \in \mathbb{N}_g\), \(S_{\Xi_N}(\hat{Z}_{\Xi_N})\) with \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\) and \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\), achieve consensus locally. According to the consensus ADMM [26], iterations for solving (P10) are expressed as

\[ \hat{Z}_{\Xi_N}^{(k+1)}:=\text{arg min}_{\hat{Z} \in \Xi_{\mathbb{P}}} \left\{ \frac{1}{\mu} \sum_{(k, i) \in \mathbb{M}_N} S_{\Xi_N}(\hat{Z}_{\Xi_N}) + \mu \left( \sum_{j \in \mathbb{N}_g} \left[ \text{Tr}(A_s^{(k+1)} C^A\left(\hat{Z}_{\Xi_N}\right)) \right] \right) + \frac{\mu}{2} \left( C^B\left(\hat{Z}_{\Xi_N}\right) - C^B\left(Z_{\Xi_N}^{(k+1)}\right) \right) \right\} \quad \forall s \in \mathbb{P} \]

where \(A_{\Xi_N}\) is the matrix vector of dual variables associated with constraints (27b) and with the same size as \(C^A\left(\hat{Z}_{\Xi_N}\right)\); in (29b), for the upper block, \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\), for the middle block, \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\), and for the lower block, \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\), \(C_{\Xi_N} \in \mathbb{K}_{\Xi_N}^k\). The objective function in (29a) is the augmented Lagrangian of (P10), and (29b) and (29c) update auxiliary variables and dual variables, respectively. We refer the readers to [26] for detailed derivations of (29).

The Frobenius norm in (29a) makes updates of primal variables not SDPs. However, by introducing slack variables, (29a) is equivalent to

\[ \hat{Z}_{\Xi_N}^{(k+1)} := \text{arg min}_{\hat{Z} \in \Xi_{\mathbb{P}}} \left\{ \sum_{j \in \mathbb{N}_g} \left[ \text{Tr}(A_s^{(k+1)} C^A\left(\hat{Z}_{\Xi_N}\right)) \right] + \frac{\mu}{2} \left( C^B\left(\hat{Z}_{\Xi_N}\right) - C^B\left(Z_{\Xi_N}^{(k+1)}\right) \right) \right\} \quad \forall s \in \mathbb{P} \]
with feasible region $\mathbb{Z}_s^\phi$ defined by
\[
\begin{bmatrix}
\varphi_s & \text{vec} \left( \zeta^A(\hat{Z}_s)-(\zeta^B(Z_s)-\frac{1}{\rho}\hat{A}_s) \right) \\
* & I
\end{bmatrix} \succeq 0.
\] (31)

Here and hereafter, "*" is used to denote partial entries of symmetric matrices. Therefore updates of primal variables can still be conducted by solving SDPs.

**Remark 4.** The size of SDP constraint (31) can be reduced by considering the symmetry of $\zeta^A(\hat{Z}_s)-(\zeta^B(Z_s)-\frac{1}{\rho}\hat{A}_s)$ and introducing multiple slack variables. For simplicity, we use $\zeta$ to represent $\zeta_s$ (the size of $\zeta$ is symmetric as long as $A_s$ is symmetric, which can be guaranteed by setting $A_0$ to a symmetric matrix vector. Furthermore, upper ($\zeta$) = lower ($\zeta$) with $\zeta$ being symmetric. Therefore, if $A_0$ is symmetric, constraint (31) can be replaced by the following equivalent form:
\[
\begin{bmatrix}
\varphi_s & \text{vec} \zeta(\hat{Z}_s) \\
* & I
\end{bmatrix} \succeq 0,
\] (32)

**B. Feasibility-Embedded Distributed Solution Approach**

Considering the SDP relaxation before being decomposed, i.e., (P8), when the relaxation is inexact, the optimal solution of (P8) is infeasible to (P7) and thus cannot be used for DID. This feasibility is determined by whether the solution satisfies rank constraints (20e). In the following, we further embed feasibility of solutions into the above solving approach by taking the rank constraints into account. Corollary 1 (see Appendix-B) gives the condition for principal submatrices in which the symmetric partially specified matrix has a completion that is not only positive semi-definite but also rank-1. With Corollary 1 and the fact that $\hat{Z}_{c_{ej}}$ contains at least one non-zero entry, rank constraints (20e) are decomposed into
\[
\text{rank}(\hat{Z}_{c_{ej}}) = 1 \quad \forall c_{ej} \in \mathcal{K} \cup \hat{\mathcal{K}}
\] (33)

Now we consider (P10) with rank constraints. First $\forall s \in \mathcal{P}$, introduce an auxiliary matrix vector variable $Y_s = [Y_{c_{ej}},Y_{\hat{c}_{ej}}]^T$ with $c_{ej} \in \mathcal{K}_s = \bigcup_{k_{ij}\in\mathcal{K}_s} (K_{ij}^L \cup K_{ij}^R)$, and $Y_s$ and $\hat{Z}_s$ are of the same size. Then (P10) with rank constraints can be formulated as
\[
\min_{\hat{Z}_s \in \mathbb{Z}_s, Y_s \in \mathbb{Y}_s} \sum_{s \in \mathcal{P}} \hat{J}_s(\hat{Z}_s)
s.t. \forall \mathcal{P}, \zeta^A(\hat{Z}_s) - \zeta^B(Z_s) = O_s
\] (34a)
\[
\hat{Z}_s - Y_s = O_{\hat{s}}
\] (34b)

The augmented (partial) Lagrangian of (34) is written as
\[
L = \sum_{s \in \mathcal{P}} L_s(\hat{Z}_s, Z_s, Y_s, A_s, \bar{A}_s)
\] (35)

with
\[
L_s(\hat{Z}_s, Z_s, Y_s, A_s, \bar{A}_s) = \hat{J}_s(\hat{Z}_s) + \text{Tr}(A_s^T(\hat{Z}_s) - \zeta^B(Z_s)) + \frac{\rho}{2}\|\zeta^A(\hat{Z}_s) - \zeta^B(Z_s)\|^2_F
\] (36)
\[
+ \text{Tr}(\hat{A}_s^T(\hat{Z}_s) - \bar{A}_s) + \frac{\bar{\rho}}{2}\|\hat{Z}_s - Y_s\|^2_F.
\]

Furthermore, iterations for solving (P11) are given by Step 1) to Step 3) as follows:

**Step 1) Update primal variables**

The update of primal variables $\hat{Z}_s$ is given as
\[
\hat{Z}_s^{(k+1)} := \arg \min_{\hat{Z}_s \in \mathbb{Z}_s} L_s(\hat{Z}_s, Z_s^{(k)}, Y_s^{(k)}, A_s^{(k)}, \bar{A}_s^{(k)}) \quad \forall s \in \mathcal{P}
\] (37)

Analogously to (30), by introducing slack variables for each subproblem, (37) is equivalent to the following SDP:
\[
\left\{ \hat{Z}_s^{(k+1)}, \cdots \right\} := \arg \min_{\{\hat{Z}_s, \varphi_s, \bar{A}_s\} \in \mathbb{Z}_s \cap \mathbb{Y}_s \cap \mathbb{Z}_s^\phi} \hat{J}_s(\hat{Z}_s) + \rho_2\varphi_s + \frac{\bar{\rho}}{2}\bar{A}_s
\] (38)

with the feasible region $\mathbb{Z}_s^\phi$ defined by (31) and $\mathbb{Z}_s^\phi$ defined by
\[
\begin{bmatrix}
\varphi_s & \text{vec}(\hat{Z}_s - (\zeta_s - \frac{1}{\rho}\hat{A}_s))^T \\
* & I
\end{bmatrix} \succeq 0.
\] (39)

**Remark 5.** If $A_0$ is symmetric, $\hat{Z}_s - (\zeta_s - \frac{1}{\rho}\hat{A}_s)$ is also symmetric (see Remark 6 for the reason). Then analogously to Remark 4, PSD constraints (39) can be reduced to multiple small PSD constraints.

**Step 2) Update auxiliary variables**

The update of auxiliary variables $Z_s$ and $Y_s$ is given as
\[
\left\{ Z_s^{(k+1)}, \cdots \right\} := \arg \min_{Z_s \in \mathcal{K}_s, \forall Y_s \in \mathbb{Y}_s} \sum_{s \in \mathcal{P}} L_s(\hat{Z}_s^{(k+1)}, Z_s, Y_s, A_s, \bar{A}_s)
\] (40)

where computing for $Z_s^{(k+1)}$ and each $Y_s^{(k+1)}$ can be conducted individually by separating (40) into
\[
Z_s^{(k+1)} = \arg \min_{Z_s \in \mathcal{K}_s} \sum_{s \in \mathcal{P}} L_s(\hat{Z}_s^{(k+1)}, Z_s, Y_s^{(k+1)}, A_s, \bar{A}_s)
\] (41a)
\[
Y_s^{(k+1)} = \arg \min_{Y_s \in \mathbb{Y}_s} \sum_{s \in \mathcal{P}} L_s(\hat{Z}_s^{(k+1)}, \hat{Z}_s, Y_s, A_s, \bar{A}_s)
\] (41b)

Here (41a) is the same as the update of auxiliary variables in consensus ADMM, which can be formulated as the simpler form given by (29b). By Proposition 3 (see Appendix-B), matrices in $Y_s$ can be updated in parallel. More importantly, updating of each matrix, i.e., $Y_{c_{ej}}$, is essentially a low rank approximation problem. This problem is non-convex due to rank constraints but an optimal solution can be given by the Eckart-Young-Mirsky Theorem [27]. Accordingly, updates of $Y_s$ can be conducted exactly and analytically as
\[
Y_s^{(k+1)} = \left[ u_{c_{ej}}^1 \ u_{c_{ej}}^2 \right]^T
\] (42)

with
\[
C_{ej} \in \mathcal{K}_s \quad \forall s \in \mathcal{P}
\] (43)

**Remark 6.** For all $\kappa \geq 0$, $\hat{Z}_s^{(k+1)}$ is symmetric, and the same for $Y_s^{(k+1)}$ according to (42). Then by (43), as long as $A_0$ is symmetric, $A_s^{(k)}$ is symmetric and thus the same for matrix
and that of $\mathcal{A}_{s}$ is given by
\[
\tilde{A}_{s}^{(\kappa + 1)} := \mathcal{A}_{s}^{(\kappa)} + \rho (\tilde{Z}_{s}^{(\kappa + 1)} - Y_{s}^{(\kappa + 1)}) \quad \forall s \in \mathcal{P}
\]

Algorithm 1: Feasibility-embedded distributed approach

Input: $N_{S}$, $\Xi_{s}$, $\rho$, $\tilde{\rho}$, $\epsilon_{abs}$ and $\epsilon_{rel}$
Output: $M$, $D$

1: Initialize $\tilde{Z}_{a}^{(0)}$, $\tilde{Y}_{a}^{(0)}$, $\tilde{A}_{a}^{(0)}$, $\tilde{A}_{s}^{(0)}$ and $\kappa \leftarrow -1$
2: repeat
3: $\kappa \leftarrow \kappa + 1$
4: for $s \leftarrow 1$ to $N_{S}$ do $\tilde{Z}_{s}^{(\kappa + 1)} \leftarrow \text{Eq. (38)}$ end for
5: $\tilde{Z}_{a}^{(\kappa + 1)} \leftarrow \text{Eq. (29b)}$
6: for $s \leftarrow 1$ to $N_{S}$ do
7: for each $C_{s,j}$ in $K_{s}$ do
8: $\{s_{k_{s,j}}^{(\kappa)}, v_{k_{s,j}}^{(\kappa)}, v_{k_{s,j}}^{(\kappa)^{T}}\} \leftarrow \text{SVD for } \tilde{Z}_{s,j}^{(\kappa + 1)} + \frac{1}{\rho} \tilde{A}_{s,j}^{(\kappa)}$
9: end for
10: $\tilde{Y}_{s}^{(\kappa + 1)} \leftarrow \text{Eq. (42)}$
11: end for
12: for $s \leftarrow 1$ to $N_{S}$ do $\{\mathcal{A}_{s}^{(\kappa + 1)}, \tilde{A}_{s}^{(\kappa + 1)}\} \leftarrow \text{Eq. (29c, 43)}$ end for
13: Compute $r^{(\kappa + 1)}$, $s^{(\kappa + 1)}$, $\epsilon_{pri}^{(\kappa + 1)}$ and $\epsilon_{dual}^{(\kappa + 1)}$
14: until $\|r^{(\kappa + 1)}\|_2 < \epsilon_{pri}^{(\kappa + 1)}$ and $\|s^{(\kappa + 1)}\|_2 < \epsilon_{dual}^{(\kappa + 1)}$
15: $m_{j} \leftarrow \tilde{Z}_{j,(1,2)}^{\text{ind}}$, $d_{j} \leftarrow \tilde{Z}_{j,(1,3)}^{\text{ind}}$, $\forall j \in N_{g}$

Finally, Algorithm 1 shows the pseudocode of the proposed feasibility-embedded distributed approach. Here computations in line 4, accounting for almost entire computational efforts, can be conducted in parallel across at most $N_{S}$ processors.

VI. CASE STUDY

The proposed numerical method for DID is tested on five systems, including the simplified 14-generator Australian (AU14Gen) [28], IEEE 14-bus, IEEE 39-bus, IEEE 118-bus and ACTIVSg200 systems [29]. Six normal steady-state operating conditions of the AU14Gen system, named case 1 to case 6 following Table 1 in [28], are also used to demonstrate the necessity of DID. IPOPT interfaced by Pyomo, and MOSEK interfaced by CVXPY, are employed to sove NLPs and SDPs, respectively. All computations are carried out on a Linux 64-Bit server with 2 Intel(R) Xeon(R) E5-2640 v4 @ 2.40GHz CPUs (a total of 40 processors provided) and 125GB RAM. Distributed computing across multiple processors for line 5 of Algorithm 1 is realized using Ray [30].

A. Parameter Setting

Two generator settings are considered to simulate different operating modes or composition of generators. For the AU14Gen system, all generators are modelled as inverters in $\mathcal{N}_{v,m}$, where both virtual inertia and damping of all inverters are dispatchable. For other test power systems, generators are set in $\mathcal{N}_{v,m}$ (or $\mathcal{N}_{d}$), $\mathcal{N}_{v}$ (low inertia), $\mathcal{N}_{d}$ (high inertia), $\mathcal{N}_{v,m}$ and $\mathcal{N}_{d}$ circularly. Parameters of each set of generators are given in Table IV in Appendix-C. Transient reactances of synchronous generators are ignored for simplicity. The base MVA is 100 MVA. For each load, $d_{l_{1}} = 0.01$ p.u. · s/rad. For each test system, $D$ contains four disturbances with parameters given in Table V of Appendix-C. Matrices $W_{1}$ to $W_{5}$ are all set to identity matrices with proper dimension. For the setting of time horizon, $t_{0} = 0$ s and $t_{f} = 30$ s. Frequency bounds $\omega^{l}(t)$ and $\omega^{h}(t)$ are give in Table VI in Appendix-C, referring to draft NEM mainland frequency operating standards of interconnected systems [17]; and $\delta = 3\pi/4$.

In the NLP formulation of DID, 3rd-order Radua collocation and $n_{k}^{l}$ = 20 are employed. In Approximation 1, $\theta_{b} = 0.580001$ to minimize the approximation error according to Appendix-A, where $\epsilon = 2.2155 \times 10^{-4}$. In the fill-reducing Cholesky factorization, $\beta_{gf} = 100$ can guarantee positive definiteness of $A_{adj} + \beta_{gf} I$ for all systems. In the feasibility-embedded distributed approach, $N_{S} = 40$, $\Xi_{s} = \{(k, 2s - 1), (k, 2s)\}$ with $k \in D$ and $s \in P$, $\epsilon_{abs} = 10^{-5}$ and $\epsilon_{rel} = 10^{-3}$ [26].

B. Numerical Results

The proposed feasibility-embedded distributed approach (FEDA) is used to solve the DID problem, for all test systems. Fig. 2 and Fig. 3 show the progress of the primal and dual residual norms by iteration, for six cases of the AU14Gen power system and other four test systems, respectively. The dashed lines show the feasibility tolerances $\epsilon_{pri}$ and $\epsilon_{dual}$. The vertical dotted lines show when the feasibility tolerance is satisfied, and the rightmost vertical dotted line shows when the stopping criterion of Algorithm 1, i.e., line 14, has been satisfied. We run the FEDA for 100 iterations to show the continued progress while the stopping criterion could be satisfied beforehand. According to Fig. 2 and Fig. 3, it can be concluded that stopping criterion of the FEDA, with certain values of penalty parameters $\rho$ and $\tilde{\rho}$, is satisfied within 100 iterations for all test systems. For the AU14Gen power system under different operating conditions, the number of iterations for convergence is slight different but all within 65 iterations. The increase of the size of power systems only causes slow growth in the number of iterations for convergence.

It can be seen that the convergence rate of $\|r\|_{2}$ and $\|s\|_{2}$ generally becomes increasingly slow as iteration progresses, which means that the FEDA can converge to modest accuracy within an acceptable number of iterations but be very slow to converge to high accuracy. This is determined by inherent convergence characteristics of ADMM [26]. Nonetheless, modest accuracy is sufficient for the practical application of DID.

It should be noted that unlike solving convex problems by ADMM where convergence can be guaranteed under mild conditions and is immune to values of penalty parameters in the augmented Lagrangian, the FEDA is not necessarily convergent. It is found that improper values of penalty parameters can lead the FEDA to be divergent, which, however, only affect convergence time when using ADMM to solve convex problems [26]. Taking the IEEE 14-bus system for example, with $(\rho, \tilde{\rho}) = (1.0, 0.5)$, the FEDA starts to drastically...
for the solution obtained by the FEDA, \( \text{rank}(\hat{\mathcal{Z}}) = 1 \) for all test systems except the IEEE 39-bus power system. But by increasing the threshold in computing \( \text{rank}(\hat{\mathcal{Z}}) \) to \( 2 \times 10^{-5} \) or taking the solution after the 62th iteration, we still have \( \text{rank}(\hat{\mathcal{Z}}) = 1 \) for the IEEE 39-bus system. Therefore, it can be concluded that for all test systems, the SDP relaxation is inexact and thus solving the decomposed SDP relaxation (P9) directly can only result in infeasible solutions to (P7). More importantly, the proposed FEDA can produce solutions which are not only with much smaller objective values than that found by NLP but also feasible to the original problem (P7) under modest tolerances. Ignoring the approximation errors between (P2) and (P7), the optimal objective value obtained by SDP gives a lower bound of objective function \( \hat{J} \) in (P2) and solutions obtained by NLP and the FEDA are both a local optimum of (P2). Clearly, the FEDA achieves a much smaller optimality gap for solutions of (P2) than NLP does for all test systems.

The effectiveness of the FEDA is also demonstrated by the time-domain results given by Fig. 4 to Fig. 7. For clarity, we call the systems with the DID results obtained by NLP and the FEDA, the NLP system and the FEDA system, respectively. Here we focus on the IEEE 14-bus system for the sake of observability, and compare time-domain curves of the NLP system and FEDA system. By Fig. 4, we can find that under the power-step disturbance, the curves of phase angle differences of branches of the NLP system and FEDA system are close, while the DID system outperforms the NLP system regarding the frequency nadir, steady-state frequency, frequency oscillation and the maximal rate of change of frequency (RoCoF). Regarding control efforts, power output changes of most generators and inverters in the FEDA system are overall larger than that in the NLP system within 0.5 s after the disturbance occurs, while in the steady state, the opposite is the case. Under the power-ramp disturbance, the DID system outperforms the NLP system regarding the steady-state frequency and the maximal RoCoF, as shown in Fig. 5. By Fig. 6, it is observed that under the power-fluctuation disturbance, fluctuations in \( \dot{\omega} \) and especially \( \omega \) of the FEDA system are smaller than that of the NLP system, while fluctuations in \( \Delta p \) of the two system are close. Under the three-phase short circuit disturbance, out-performance of the FEDA system is more

### Table III

| Test systems     | Optimal objective values | rank(\(\hat{\mathcal{Z}}\))
|------------------|--------------------------|-----------------------|
| NLP              | SDP                      | FEDA                  | SDP | FEDA  |
| Case 1           | 187.95                   | 93.67                 | 127.44 | 21     |
| Case 2           | 265.61                   | 120.66                | 191.14 | 15     |
| Case 3           | 106.74                   | 42.85                 | 74.95  | 14     |
| Case 4           | 61.547                   | 39.67                 | 46.10  | 20     |
| Case 5           | 108.05                   | 59.14                 | 83.15  | 21     |
| Case 6           | 71.123                   | 37.86                 | 57.67  | 20     |
| IEEE 14-bus      | 116.86                   | 32.55                 | 54.98  | 12     |
| IEEE 39-bus      | 269.99                   | 152.25                | 183.09 | 17     |
| IEEE 118-bus     | 100.21                   | 63.93                 | 88.72  | 12     |
| ACTIVSg200       | 115.47                   | 70.40                 | 81.98  | 7      |

* Threshold is set to \( 10^{-5} \) below which eigenvalues are considered zero.
significant, as shown in Fig. 7. Except for the out-performance regarding the frequency nadir, frequency oscillation and the maximal RoCoF, the FEDA system also has overall smaller oscillations in phase angle differences of branches and power output changes of generators or inverters.

Fig. 8 compares computation time of NLP, SDP and the FEDA spent on solving DID of each test system. The use of open-source modeling languages, i.e., Pyomo and CVXPY, causes significant time spent on constructing and passing models, which, however, can be reduced to a negligible amount by employing C++ interface of IPOPT and Fusion interface of MOSEK. Thus computation time of NLP and SDP only includes the time spent by optimizers, and time spent on line 4 of Algorithm 1 equals to the maximal time spent by MOSEK in each processor. In Fig. 8, we can see that NLP has a computation time advantage among the three approaches, being undermined as the size of power systems increases. As the cost of achieving solutions with a smaller optimality gap, the FEDA is inevitably with the maximum computation time for most cases, which, however, is acceptable profiting from distributed parallel computing. For the six cases of AU14Gen system, computation time of the FEDA is even very close or less than that of SDP.

Fig. 9 shows dispatch results of virtual inertia and damping for the AU14Gen system under different operating conditions, obtained by the FEDA. We can see that following variation in operating conditions, most generators need to significantly adjust their virtual inertia, damping or both of them to optimize the system performances and control efforts. This, to some extend, demonstrates the necessity of DID for operation of future power grids with high heterogeneity in operating conditions, to ensure a optimal tradeoff between synchronism performances, frequency performances and control efforts.

VII. CONCLUSION

This paper numerically addresses the DID problem for future inverter-dominant transmission networks. By the Radua collocation method, the DID problem is first formulated as a NLP with flexibility to handling time-varying performance constraints and various types of disturbances. Next, the highly non-convex NLP is relaxed into a convex SDP for which sparsity is exploited to improve computational efficiency. Finally, a feasibility-embedded distributed solution approach is proposed under the framework of ADMM. Numerical experiments on five test systems demonstrate that by tuning penalty parameters, the proposed solution approach can converge to modest accuracy within several tens of iterations as well as acceptable computation time benefiting from distributed parallel computing. The SDP relaxation of the NLP of DID is inexact while the feasibility-embedded distributed approach can produce solutions being not only feasible to the original problem but also with much smaller optimality gaps than that achieved by the local solution approach. Variations in dispatch results under different operating conditions for the AU14Gen system demonstrates the necessity of DID for power grids with increasingly high heterogeneity in operating conditions.

For the future direction, firstly, influence of penalty parameters on convergence of the feasibility-embedded distributed approach will be further investigated to develop tuning or adaptive adjustment schemes; secondly, impacts of fidelity of system dynamic models on results of DID will be evaluated for determining appropriate model fidelity that balances the computational complexity and accuracy for the DID problem,
Fig. 8. Comparison of computation time of NLP, SDP, and the FEDA.

Fig. 9. Dispatch results of virtual inertia and damping for the AU14Gen system under operating condition case 1 to case 6, obtained by the FEDA.

and model reduction for parts of system more remote from the fault can potentially deal with the case involving high-fidelity models; and thirdly, distributed DID independent of central control centers is worth pursuing, where learning-based approaches can promisingly tackle complications caused by possible divergence of distributed algorithms for non-convex problems and potential real-time execution of DID.

ACKNOWLEDGMENT

The authors would like to thank Dr. Xiao Han, from Microsoft Research Asia, for the support of Linux servers.

REFERENCES

[1] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbič, “Foundations and challenges of low-inertia systems,” in 2018 Power Systems Computation Conference (PSCC). IEEE, 2018, pp. 1–25.

[2] A. Ulbig, T. S. Borsche, and G. Andersson, “Impact of low rotational inertia on power system stability and operation,” IFAC Proceedings Volumes, vol. 47, no. 3, pp. 7290–7297, 2014.

[3] H. Bevrani, T. Ise, and Y. Miura, “Virtual synchronous generators: A survey and new perspectives,” International Journal of Electrical Power & Energy Systems, vol. 54, pp. 244–254, 2014.

[4] M. Garmroodi, G. Verbič, and D. J. Hill, “Frequency support from wind turbine generators with a time-variable droop characteristic,” IEEE Trans. Sustain. Energy, vol. 9, no. 2, pp. 676–684, 2018.

[5] E. Mallada, “idroop: A dynamic droop controller to decouple power grid’s steady-state and dynamic performance,” in 2016 IEEE 55th Conference on Decision and Control (CDC). IEEE, 2016, pp. 4957–4964.

[6] Y. Jiang, R. Pates, and E. Mallada, “Dynamic droop control in low-inertia power systems.”

[7] P. Jacquod and L. Pagnier, “Optimal placement of inertia and primary control in high voltage power grids,” in 2019 53rd Annual Conference on Information Sciences and Systems (CISS). IEEE, 2019, pp. 1–6.

[8] L. Pagnier and P. Jacquod, “Optimal placement of inertia and primary control: a matrix perturbation theory approach,” arXiv preprint arXiv:1906.06922, 2019.

[9] B. K. Poolla, D. Gros, and F. Dörfler, “Placement and implementation of grid-forming and grid-following virtual inertia and fast frequency response,” IEEE Transactions on Power Systems, vol. 34, no. 4, pp. 3535–3546, jul 2019.

[10] A. Ademola-Idowu and B. Zhang, “Optimal design of virtual inertia and damping coefficients for virtual synchronous machines,” arXiv preprint arXiv:1806.08488, 2018.

[11] D. Groß, S. Bolognani, B. K. Poolla, and F. Dörfler, “Increasing the resilience of low-inertia power systems by virtual inertia and damping,” in IREP Bulk Power System Dynamics & Control Symposium, 2017.

[12] T. Borsche and F. Dörfler, “On placement of synthetic inertia with explicit time-domain constraints,” arXiv preprint arXiv:1705.03244, 2017.

[13] B. K. Poolla, S. Bolognani, and F. Dörfler, “Optimal placement of virtual inertia in power grids,” IEEE Trans. Autom. Control, vol. 62, no. 12, pp. 6209–6220, 2017.

[14] T. S. Borsche, T. Liu, and D. J. Hill, “Effects of rotational inertia on power system damping and frequency transients,” in Decision and Control (CDC), 2015 IEEE 54th Annual Conference on. IEEE, 2015, pp. 5940–5946.

[15] S. Yan, “Controlling time-variant virtual inertia from storage by dynamic programming and propt,” arXiv preprint arXiv:1903.03790, 2019.

[16] A. R. Bergen and D. J. Hill, “A structure preserving model for power system stability analysis,” IEEE Trans. Power App. Syst., vol. PAS-100, no. 1, pp. 25–35, Jan 1981.

[17] R. Panel, “Review of the frequency operating standard – stage two, final determination,” 2019.

[18] L. T. Biegler, Nonlinear programming: concepts, algorithms, and applications to chemical processes. Siam, 2010, vol. 10.

[19] A. U. Raghunathan and A. V. Knyazev, “Degeneracy in maximal clique decomposition for semidefinite programs,” in 2016 American Control Conference (ACC). IEEE, 2016, pp. 5605–5611.

[20] Y. Zheng, G. Fantuzzi, and A. Papachristodoulou, “Decomposition methods for large-scale semidefinite programs with choral aggregate sparsity and partial orthogonality,” in Large-Scale and Distributed Optimization. Springer, 2018, pp. 33–55.

[21] L. Vandenberghe, M. S. Andersen et al., “Chordal graphs and semidefinite optimization,” Foundations and Trends® in Optimization, vol. 1, no. 4, pp. 241–433, 2015.

[22] F. Cazals and C. Karande, “A note on the problem of reporting maximal cliques,” Theoretical Computer Science, vol. 407, no. 1-3, pp. 564–568, 2008.

[23] J. L. Gross and J. Yellen, Graph theory and its applications. Chapman and Hall/CRC, 2005.

[24] D. K. Molzahn, J. T. Holzer, B. C. Lesieutre, and C. L. DeMarco, “Approximate initialization for large-scale semidefinite programming,” in 2013 IEEE 52nd Annual Conference on Decision and Control (CDC). IEEE, 2013, pp. 982–987.

[25] Y. Zhang, “Restricted low-rank approximation via admm,” arXiv preprint arXiv:1512.01748, 2015.

[26] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein et al., “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Foundations and Trends® in Machine learning, vol. 3, no. 1, pp. 1–122, 2011.

[27] C. Eckart and G. Young, “The approximation of one matrix by another of lower rank,” Psychometrika, vol. 1, no. 3, pp. 211–218, 1936.

[28] M. Gibbard and D. Wollves, “Simplified 4-generator model of the se australian power system. 2008,” The University of Adelaide, South Australia.

[29] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, “Grid structural characteristics as validation criteria for synthetic networks,” IEEE Trans. Power Syst., vol. 32, no. 4, pp. 3258–3265, 2016.

[30] M. Moritz, R. Nishihara, S. Wang, A. Tumanov, R. Liaw, E. Liang, M. Elibol, Z. Yang, W. Paul, M. I. Jordan et al., “Implementation of multipliers,” Foundations and Trends® in Optimization, vol. 1, no. 2, pp. 33–55, 2016.

[31] P. Moritz, R. Nishihara, S. Wang, A. Tumanov, R. Liaw, E. Liang, M. Elibol, Z. Yang, W. Paul, M. I. Jordan et al., “Implementation of multipliers,” Foundations and Trends® in Optimization, vol. 1, no. 2, pp. 33–55, 2016.

[32] J. Dancis, “Positive semidefinite completions of partial hermitian matrices,” Linear Algebra Appl., vol. 175, pp. 97–114, 1992.
### APPENDIX

#### A. Numerical Analysis of Approximation 2

The essence of Approximation 1 is using quadratic function $s_1$, linear function $s_2$ and quadratic function $s_3$ to approximate $\sin \theta$ for $\theta \in [-\theta_c, -\theta_b]$, $\theta \in [-\theta_b, \theta_b]$ and $\theta \in [\theta_b, \theta_c]$, respectively. Clearly, $\theta_b$ can observably impact approximation errors and should be selected carefully. Define approximation error function $e : \theta_b \rightarrow t_{\theta_b}^0 (\sin \theta - \beta^T \hat{c})^2 d \theta$ and function $t_{\theta_b}^0 : \theta_c \rightarrow \{ \arg \min_{\theta_b} e, \text{s.t.} \theta_b \in [0, \frac{\pi}{2}] \}$. Graphs of $\theta_b - \theta_c, \theta_c - t_{\theta_b}^0$ and $\theta_b - \min_{\theta_b} e$ are shown in Fig. 10. Numerically, we can find that $\forall \theta_b \in [\frac{\pi}{2}, \pi]$, $e$ is a convex function in domain $[0, \frac{\pi}{2}]$. Thus, given $\theta_b \in [\frac{\pi}{2}, \pi]$, by solving $\{ \arg \min_{\theta_b} e, \text{s.t.} \theta_b \in [0, \frac{\pi}{2}] \}$, the unique optimal value of $\theta_b = t_{\theta_b}^0$ to minimize the approximation error can be obtained. Additionally, we can also find that for $\theta_b \in [\frac{\pi}{2}, \pi]$, Approximation 1 can be with very high accuracy.

![Fig. 10. Numerical analysis of Approximation 2.](image)

#### B. Propositions and Corollaries

**Proposition 1.** For a given graph $G(V, E)$, define a graph $\hat{G}(\hat{V}, \hat{E})$ where each node in $\hat{V}$ represents a set of nodes in $V$: $\bigcup_{V \in \hat{V}} \bigcup_{V' \in V} \hat{V} = V$; $\forall \hat{V}, \hat{V}' \in \hat{V}, \hat{V} \neq \hat{V}' \Rightarrow \hat{V} \cap \hat{V}' = \emptyset$; and $(\hat{i}, \hat{j}) \in \hat{E}$ satisfies $\exists (i, j) \in E$. Let $K_l = \{ C_1, ..., C_{nmc} \}$ and $\hat{K}_l = \{ \hat{C}_1, ..., \hat{C}_{nmc} \}$ be sets of all maximal cliques of $G$ and $\hat{G}$, respectively. Then the following statements hold:

(i) $G(V, E)$ is chordal iff $G(\hat{V}, \hat{E})$ is chordal and $\forall i \in \{1, ..., n_{mc}\}$, the induced subgraph of $G(\hat{V}, \hat{E})$ by the node set $\hat{C}_i$ is chordal; and

(ii) denote by $K_l$ the set of all maximal cliques of the induced subgraph of $G(V, E)$ by the node set $C_i$, then $K = \bigcup_{i \in \{1, ..., n_{mc}\}} K_i$.

**Proof.** The proof is trivial with basic properties of graphs and the definition of maximal cliques.

**Proposition 2.** Denote by $G_n(N, B)$ the underlying graph of the power network, then we have

(i) graph $G$ is chordal if $G_n$ is chordal;

(ii) assume that $G_n$ is chordal with its set of maximal cliques denoted by $K_n = \{ C_{nj} \}$ since we can always find a chordal extension for $G_n$. Maximal cliques of $G$ are given by $K = \bigcup_{K_n \in D \in C_{nj}} K^k_n$ where $K^k = K^k_{\text{nc}} \cup K^k_{\text{d}} \cup K^k_{\text{e}}$ with $K^k_{\text{nc}} = \{ C_{nj} | C_{nj} \subseteq (i \text{, } j \in \{1, ..., n_{mc}\}) \};$ $K^k_{\text{d}} = \{ C_{nj} | C_{nj} \subseteq (i \text{, } j \in \{1, ..., n_{mc}\}) \};$ $K^k_{\text{e}} = \{ C_{nj} | C_{nj} \subseteq (i \text{, } j \in \{1, ..., n_{mc}\}) \}$.

**Proof.** Fig. 11(a) shows the aggregate sparsity pattern of matrix $Z$ at the block level with $Z$ broken into blocks corresponding to each pair of disturbances and time elements.

**Corollary 1.** Let $G(V, E)$ be a chordal graph and let $K =
{C_1, ..., C_{n_{mc}}} be the set of all maximal cliques. Then \( Z' \in \mathbb{S}^{\mathcal{N}} \) has a rank-1 positive semi-definite completion if and only if \( Z' \) satisfies that for each \( i \in \{1, ..., n_{mc}\}, \mathcal{S}_C'(Z') \succeq 0 \) and max\{rank(\( \mathcal{S}_C(Z') \))\} = 1.

Proof. Corollary 1 is a corollary of [31, Therom 2.5] and [32, Theorem 1.5].

Proposition 3. Equation (41b) is equivalent to

\[
Y_s^{(k+1)} = \left[ \arg \min \left\{ \|Y_{c,j} - \left( \hat{Z}_s^{(k+1)} + \frac{1}{\rho} A_s^{(k)} \right) \| \right\} \frac{\gamma}{\rho} \right]^T
\]

with \( C_{c,j} \in K_s \quad \forall s \in P \)

Proof. In (41b), the objective function is given by

\[
L_s(\hat{Z}_s^{(k+1)}; \mathcal{Z}_s^{(k)} Y_s, A_s^{(k)}, \hat{A}_s^{(k)}) = \frac{\rho}{2} \|Y_s - \hat{Z}_s^{(k+1)} - \frac{1}{\rho} A_s^{(k)} \|^2_F
\]

\[
- \frac{1}{2\rho} \| \hat{A}_s^{(k)} \| + \text{Tr}(\hat{A}_s^{(k)} Y_s) + \frac{\rho}{2} \| \mathcal{Z}_s^{(k)} - \hat{Z}_s^{(k+1)} \|_F^2
\]

and \( Y_s \) is involved only in the first term. Dropping other terms and \( \rho/2 \) in the first term results in (47).

C. Parameter Settings

| TABLE IV | PARAMETERS OF GENERATORS. |
|----------|--------------------------|
| Generator | \( m_1 \) | \( m_4 \) | \( d_i \) | \( d_j \) |
| \( N_{m} \) and \( N_{g} \) | 0.5\( m_1 \) | 0.5\( m_4 \) | 0.5\( d_i \) | 0.5\( d_j \) |
| \( N_{g} \) (low inertia) | 0.01\( m_1 \) | 0.01\( m_4 \) | 0.01\( d_i \) | \( d_j \) |
| \( N_{g} \) (high inertia) | 0.5\( m_1 \) | 0.5\( m_4 \) | 0.01\( d_i \) | \( d_j \) |
| \( N_{g} \) sym | 0.01\( m_1 \) | \( m_4 \) | 0.5\( d_i \) | 0.5\( d_j \) |
| \( N_{g} \) (low inertia) | 0.01\( m_1 \) | \( m_4 \) | 0.01\( d_i \) | \( d_j \) |

Note: \( \hat{m}_i = \frac{10p_{g, \text{max}}}{\omega_{\text{syst}}} \) with \( p_{g, \text{max}} \) being maximal steady-state active power output of generator \( i \) and \( \omega_{\text{syst}} \) being the synchronous angular speed. \( d_i = \frac{2p_{g, \text{max}}}{2\pi} \). \( p_0 = -\frac{\rho}{\rho_0} = 3p_{g, \text{max}} \) for each generator.

| TABLE V | FREQUENCY BOUNDS. |
|----------|------------------|
| Disturbance | Time interval | \( \omega^b(t) \) | \( \omega^s(t) \) |
| \( D_1 \) and \( D_2 \) | \([0s, 15s] \cup [15s, 30s]\) | \([49.5, 49.85]\) | \([50.5, 50.15]\) |
| \( D_3 \) | \([0s, 30s]\) | 49.85 | 50.15 |
| \( D_4 \) | \([0s, 15s] \cup [15s, 30s]\) | \([49, 49.5]\) | \([51, 50.5]\) |

Note: The above table shows location of disturbances, with numbers denoting bus number. Test systems are at the equilibrium point at \( t = t_0 \), each disturbance occurs at \( t = t_0 \), and \( P_0 \) denotes the initial load power or generation power. For disturbances in \( D_1 \), step amplitude is set to \(-50\% P_0 \), where \( P_0 \) denotes the initial load power for load buses or generation power for generator buses. For disturbances in \( D_2 \), height of ramp and duration of ramp are set to \(-50\% P_0 \) and 5 s, respectively. Disturbances in \( D_3 \) are emulated by a random power disturbance which changes its value randomly at a equal interval being 0.5 s according to a uniform distribution with the interval being \([-20\% P_0, 20\% P_0]\). Disturbances in \( D_4 \) are assumed occurring at the middle of the branch, with short circuit resistance being 0, and being cleared by disconnecting the two sides breakers of the branch after 0.1 s. \( \rho_0 \) is set to 0.15, 0.15, 0.6 and 0.1 for disturbances in \( D_1, D_2, D_3 \) and \( D_4 \), respectively.