Sparsity-Constrained Games and Distributed Optimization with Applications to Wide-Area Control of Power Systems

Feier Lian, Alexandra Duel-Hallen Fellow, IEEE, and Aranya Chakrabortty Senior Member, IEEE

Abstract—Multi-agent networked dynamic systems attracted attention of researchers due to their ability to model spatially separated or self-motivated agents. In particular, differential games are often employed to accommodate different optimization objectives of the agents, but their optimal solutions require dense feedback structures, which result in high costs of the underlying communication network. In this work, social linear-quadratic-regulator (LQR) optimization and differential games are developed under a constraint on the number of feedback links among the system nodes. First, a centralized optimization method that employs the Gradient Support Pursuit (GraSP) algorithm and a restricted Newton step was designed. Next, these methods are combined with an iterative gradient descent approach to determine a Nash Equilibrium (NE) of a linear-quadratic game where each player optimizes its own LQR objective under a shared global sparsity constraint. The proposed noncooperative game is solved in a distributed fashion with limited information exchange. Finally, a distributed social optimization method is developed. The proposed algorithms are used to design a sparse wide-area control (WAC) network among the sensors and controllers of a multi-area power system and to allocate the costs of this network among the power companies. The proposed algorithms are analyzed for the Australian 50-bus 4-area power system example.

Index Terms—Game Theory, Nash Equilibrium, Sparsity, Distributed Optimization, Power Networks

I. INTRODUCTION

Multi-agent networked dynamic systems model many practical scenarios where the agents are spatially separated or have different economic priorities, e.g., cyber-physical power networks, multi-vehicle formation, intelligent transportation, industrial automation, etc [1]. When the agents have different optimization objectives, the resulting optimal control problem can be modeled as a noncooperative differential game [2] where the agents are referred to as players, and the players’ control vectors represent their strategies. In particular, linear-quadratic differential games received significant attention in the literature due to their tractability, feasibility of distributed implementation, and broad applicability of the quadratic utility function [2], [5].

The traditional state-feedback centralized LQR optimization [3] and the linear-quadratic games [2], [4], [5] require a dense feedback matrix from all states to all controllers, thus resulting in significant investment into the resulting communication network. For state-feedback optimal-control LQR, algorithms for sparsity promotion were addressed in [6], and for satisfying structural constraints in the feedback matrix in [7], [8], respectively. More generally, sparsity-constrained optimization has been investigated in [9], [10]. However, these methods employ global optimization objectives and centralized implementation, which limit their applicability to multi-agent systems. Moreover, distributed approaches, such as optimal control with sparse state feedback [11] and the Distributed Alternating Direction Method of Multipliers (D-ADMM) [12], do not accommodate flexible sparsity constraints and different optimization objectives of multiple agents.

In this paper, we investigate sparsity-constrained LQR optimization for dynamic systems with linear state feedback. We impose a constraint expressed by the off-diagonal cardinality of the feedback matrix on the maximum number of communication links that transmit state-feedback data among different system nodes. First, to solve the centralized LQR optimal control problem under an off-diagonal-cardinality constraint, we employ the greedy Gradient Support Pursuit (GraSP) algorithm [9], which was shown to provide accurate approximations to sparsity-constrained solutions for a wide class of optimization functions. The proposed method also utilizes the restricted Newton step [7] to speed up convergence. Second, we develop solutions for linear-quadratic games under global communication constraints. To compute the noncooperative equilibria of these games, we combine the ideas of GraSP and iterative gradient descent approaches [13], [14]. In the resulting algorithm, the computation of the players’ utilities is distributed and requires limited information exchange. To retain these features when performing the social optimization, we convert the noncooperative game into a potential game [14] where the players’ utilities agree, thus producing a social sparsity-constrained distributed algorithm.

The proposed algorithms are applied to the wide-area control (WAC) of power systems, which helps in suppression of inter-area oscillations, but potentially requires a substantial investment into the communication network needed to exchange state feedback among the power companies [15]–[17]. To model WAC, the cyber-physical power network is viewed as a multi-agent system, where the agents represent different power companies. In addition to designing cost-efficient wide-area controllers, we apply the proposed sparsity-constrained algorithms to allocate the costs of WAC among the power companies, which impacts economic adoption, fea-
sibility, fairness, and customer billing for emerging smart grid systems [15], [18]. The cost-allocation problem is formulated using a cooperative network-formation game (NFG), with power companies as players. The proposed distributed cost allocation method characterizes each company’s need for WAC in terms of the energy it saves by utilizing feedback and cooperation over a range of sparsity levels. It improves on our previous WAC cost allocation approaches [17], [19], which employed heuristics, relied on the centralized optimization in [6], and extrapolated the costs of a dense network [5] to sparse scenarios.

The main contributions of this paper are:

- Development and analysis of centralized and distributed social optimization algorithms for multi-agent sparsity-constrained LQR optimal control with static linear state feedback;
- Design and validation of a noncooperative linear-quadratic control input of node
- the vector of control inputs, Whenever
- $K$ represents the feedback gain from state
- $x$ is a scalar impulse
- $u(t)$ are organized according to their physical locations, the matrix $K$ is in the form

$$K = \begin{bmatrix}
K_{11} & K_{12} & \cdots & K_{1n} \\
K_{21} & K_{22} & \cdots & K_{2n} \\n\vdots & \vdots & \ddots & \vdots \\
K_{n1} & K_{n2} & \cdots & K_{nn}
\end{bmatrix}$$

(3)

where the block $K_{ij} \in \mathbb{R}^{m_i \times m_j}$ represents feedback of the states of node $j$ to the controllers of node $i$. The diagonal block $K_{jj}$ corresponds to the local feedback links within node $j$, and we assume that these links do not incur communication costs. Without loss of generality, we define the communication cost as the number of communication links associated with the off-diagonal blocks of $K$, given by

$$\text{card}_{\text{off}}(K) = \sum_{i,j=1,i\neq j}^{n} \text{nnz}(K_{ij})$$

(4)

where $\text{nnz}(\cdot)$ operator counts the number of nonzero elements in a matrix. The proposed algorithms can be easily adapted to other sparsity criteria. Additional notation used in the proposed algorithms is described in Table I.

In this section, we solve the social LQR problem with full state feedback, subject to a constraint on the number of communication links. The social global LQR objective function is given by [20]

$$J(K) = \int_{t=0}^{\infty} |x(t)^T Q x(t) + u(t)^T R u(t)| dt$$

(5)

Under the sparsity constraint $s$, the global LQR problem is

| Term | Definition |
|------|-------------|
| $||K||_F$ | Frobenius norm of the matrix $K$, defined by $\text{trace}(K^T K)$. |
| $\text{supp}(K)$ | The support set of the matrix $K$, i.e., the set of indices of the nonzero entries of matrix $K$ [9]. |
| $[K]_s$ | The matrix obtained by preserving only the $s$ largest-magnitude entries of the matrix $K$, and setting all other entries to zero. |
| $K^{\text{off}}$ | The matrix obtained by preserving only the off-diagonal blocks of the matrix $K$ (see (3)) and setting all other entries to zero. |
| $K^{\text{diag}}$ | The matrix obtained by preserving the diagonal blocks of the matrix $K$ and setting all other entries to zero. |
| $\nabla_K J(K)$ | The gradient of the scalar function $J(K)$ with respect to the matrix $K$ [20]. Assuming $K \in \mathbb{R}^{m \times n}$, $\nabla_K J(K)$ is given by a $m \times n$ matrix with the elements $[\nabla_K J(K)]_{ij} = \partial J(\mathbf{x}(t)) / \partial x_{ij}$, |
| $\nabla_K J(K)|_T$ | The gradient of the scalar function $J(K)$ with respect to the matrix $K$ projected onto the index set $T$. The matrix $\nabla_K J(K)|_T$ is obtained by preserving only the entries of $\nabla_K J(K)$ with indices in the set $T$ and setting other entries to zero. |
| $\Delta_{\text{AV}}(x, T)$ | The restricted Newton step vector $\Delta_{\text{AV}}(x; T)$ at matrix $K \in \mathbb{R}^{m \times n}$ under the structural constraint $\text{supp}(K) \subseteq T$. First, the $m \times 1$ vector $x$ is computed by stacking the columns of $K$, and the function $g(x)$ is defined as $g(x) \triangleq f(x)$. Then the $m \times 1$ restricted Newton step vector $\Delta_{\text{AV}}(x; T)$ of $g(x)$ at $x$ [9] is computed using the conjugate gradient (CG) method [11]. The vector $\Delta_{\text{AV}}(x, T)$ is then converted into an $m \times n$ matrix by stacking the consecutive $m \times 1$ segments of $\Delta_{\text{AV}}(x; T)$. |
formulated for system (12) as
\[
\min_K J(K)
\]
\[
\text{s.t. } \text{card}_{\text{off}}(K) \leq s
\]
\[
\dot{x}(t) = Ax(t) + Bu(t) + Dw(t)
\]
\[
u(t) = -Kx(t)
\]
where \(Q \in \mathbb{R}^{m \times m} \geq 0, R \in \mathbb{R}^{q \times q} > 0\) are the design matrices. Direct solution of (6) can have combinatorial complexity \([9]\). We utilize the Gradient Support Pursuit (GraSP) method \([5]\) in the proposed centralized method in Algorithm 1. For \(s = 1\), it is initialized with a feasible zero- or low-cost matrix as discussed in Section IV-C. If optimization for a range of sparsity constraints is desired, the algorithm is performed in the order of increasing \(s\)-values and is initialized with the solution of the previous optimization. Given the overall sparsity constraint \(s\), in Steps 2(1–4) of each iteration, the algorithm extends the matrix \(K\) along its steepest 2s gradient-descent directions. In Step 2(4), the matrix \(K\) is updated using the restricted Newton step, and the step size \(\lambda\) is chosen via the Armijo line search \([22]\). In Step 2(5), pruning is performed to impose the constraint \(s\). To guarantee stability of the feedback matrix after pruning, we provide a backtracking option to return to a previously found stable solution, which has \(s - 1\) or fewer communication links. Note that the stopping criteria in Step 2(6) are also used to determine convergence in the sparsity-promotion algorithm \([6]\).

B. Sparsity-Constrained Linear-Quadratic Games and Distributed Optimization

Suppose that there are \(r\) agents, where the agent \(i\) owns \(n_i\) nodes in the system (1). Without loss of generality, the nodes are partitioned as follows
\[
S_1 = \{1, 2, \ldots, n_1\} \Rightarrow \text{belongs to agent 1.}
\]
\[
S_2 = \{n_1 + 1, n_1 + 2, \ldots, n_1 + n_2\} \Rightarrow \text{belongs to agent 2.}
\]
\[
\ldots \ldots \ldots \ldots
\]
\[
S_r = \{n_1 + n_2 + \ldots + n_{r-1} + 1, \ldots, n\} \Rightarrow \text{belongs to agent } r.
\]

We can rewrite the states and control inputs of each agent \(i\) in (1) as follows:
\[
\dot{x}_i(t) = \sum_{k=1}^{r} A_{ik}x_k(t) + \sum_{k=1}^{r} B_{ik}u_k(t) + D_{ik}w(t)
\]
\[
u_i(t) = -K^i x_i(t)
\]
where \(x_i(t) = (x_{n_1+1, n_1+2, \ldots, n_1+n_2, n_1+1+n_2, \ldots, n_1+n_2+1})^T \in \mathbb{R}^{M_i \times 1}\) is the vector of states for agent \(i, M_i = \sum_{j=n_1+1}^{n_1+n_2+1} m_j;\)
\[
u_i(t) = (u_{n_1+1, n_1+2, \ldots, n_1+n_2, n_1+1+n_2, \ldots, n_1+n_2+1})^T \in \mathbb{R}^{N_i \times 1}\]
\(K^i\) is the submatrix of \(K\) associated with the control inputs of the agent \(i, D_i\) is the control matrix for the disturbance input that enters agent \(i,\) and \(D = \text{col}(D_1, \ldots, D_r).\)

Next, we briefly summarize results on linear-quadratic games for this system \([2, 5]\). The agents in (7) are viewed as players that optimize their individual objectives \(J_i\) by selecting their control inputs \(u_i(t),\) for \(i = 1, \ldots, r.\) The objective of the player \(i\) is given by
\[
J_i(u_1, u_2, \ldots, u_r) = \int_0^\infty [x(t)Q_i x(t)^T + u_i(t)^T R_i u_i(t)] dt
\]
where \(R_i > 0 \in \mathbb{R}^{N_i \times N_i},\) and \(Q_i \in \mathbb{R}^{m \times m}\). A Nash Equilibrium (NE) is achieved when it is impossible for any player to improve its objective function by unilaterally changing its strategy. At a given NE, the players employ Nash strategies \((u_1^*(t), u_2^*(t), \ldots, u_r^*(t))\) defined as
\[
J_i(u_1^*(t), u_2^*(t), \ldots, u_r^*(t)) \leq J_i(u_1(t), u_2^*(t), \ldots, u_r^*(t)), \forall u_i(t), t \in [0, \infty) (11)
\]
for \(i \in \{1, r\}\), where \(u_{-i}(t) := (u_1(t), \ldots, u_{i-1}(t), u_{i+1}(t), \ldots, u_r(t))\) is the tuple of strategies formed by all players except for the player \(i.\) When state feedback is employed \([9]\), the Nash strategies \(u_i^*(t)\) in eq. (11) can be determined by solving the cross-coupled algebraic Riccati equations (CARE) (eq.(8) in \([3]\)). In \([5]\), it is shown that there exists a unique solution to CARE under the assumption of sufficiently weak coupling among the players. An iterative algorithm provided in \([3]\) for finding this solution produces a dense feedback matrix since CARE does not impose sparsity constraints.

Next, we formulate a linear-quadratic game where each player minimizes its selfish LQR objective under a shared global communication link constraint \(s.\) Thus, Nash strategies \((u_1^*(t), u_2^*(t), \ldots, u_r^*(t))\) satisfy for \(i \in \{1, \ldots, r\}\)
\[
J_i(u_1^*(t), u_2^*(t), \ldots, u_r^*(t)) \leq J_i(u_1(t), u_2^*(t), \ldots, u_r^*(t)), \forall u_i(t) (12)
\]
Equivalently, since linear state feedback \([9]\) is employed, the

Algorithm 1 Minimizing the centralized LQR objective under the off-diagonal sparsity constraint \(s.\)

1. Initialization
\(K := K_0\)

2. Iteration
while stopping criteria not met do
\(K^{prev} := K\)
(0) Compute gradient of \(J(K)\) w.r.t \(K: g = \nabla_K J(K)\)
(1) Identify up to 2s off-diagonal block directions: \(Z = \text{supp}(\{g^{\text{off}}\}_{2s})\)
(2) Merge support: \(T = Z \cup \text{supp}(K)\)
(3) Descend using the Newton step of \(J\) restricted to \(T:\)
\(K := K + \lambda \Delta_{\text{new}}(K, T)\)
(5) Prune communication links: \(K := K^{\text{diag}} + [K^{\text{off}}]_s\)
(6) Stopping criteria:
\(|K - K^{prev}| \leq \epsilon_{\text{abs}} \sqrt{q m} + \epsilon_{\text{rel}} ||K^{prev}||_2.\)
end while

3. Polishing
\(T = \text{supp}(K)\)
while not \(|\nabla_K J| z \leq \epsilon_2\) do
Descend using the Newton step of \(J\) restricted to \(T:\)
\(K := K + \lambda \Delta_{\text{new}}(K, T)\)
end while
strategy of player $i$ is given by the submatrix $K^i$, and Nash strategies are expressed as $(K^{i*}, K^{-i*})$, which satisfy for each $i \in \{1, ..., r\}$

$$ J_i(K^{i*}, K^{-i*}) \leq J_i(K^i, K^{-i*}), \forall K^i $$

s.t. $\text{card}_{off}(K^i) \leq s$ (13)

where the tuple $K^{-i} := (K^1, ..., K^{i-1}, K^{i+1}, ..., K^r)$ represents the strategies of all other players except $i$.

The proposed game is described in Algorithm 2. It is inspired by the iterative gradient descent methods in [13], [14], where each player takes a small step towards minimizing its own objective while other players’ strategies are fixed. In each step associated with player $i$, we use the GraSP algorithm [9] to update the strategic variable $K^i$ while maintaining the overall sparsity constraint. Thus, in the submatrix $K^i$ (9), the elements representing local feedback, i.e., those in the blocks $K_{jj} \in \mathbb{R}^{s_i \times m}$ in (3) for $j \in S_i$, are free variables while the off-diagonal blocks in (2) are subject to the sparsity constraint. The effect of different initial link allocations is discussed in Section IV.

Algorithm 2 is distributed in the sense that player $i$ individually updates its strategic variable $K^i$. Each player has prior knowledge of the underlying physical system (i.e., the matrices $A, B, D$ in (1)) and broadcasts its strategic move (the submatrix $K^i$) after its strategy is updated at the completion of Step 2(6) and each inner loop of Step 3.

Finally, note that the social optimization under the sparsity constraint (6) can also be implemented in a distributed fashion using a potential game [14], where the individual objectives in (13) are replaced with the common social utility (5) while the players’ strategies are still defined as their control vectors. The equilibria of this game, which we refer to as the distributed social optimization algorithm, are defined as

$$ J(K^{i*}, K^{-i*}) \leq J(K^i, K^{-i*}), \forall K^i $$

s.t. $\text{card}_{off}(K^i) \leq s$ (14)

These equilibria can be computed using Algorithm 2, where the social objective $J()$ replaces $J_i()$ in Step 2(2) and in the polishing step.

### III. Example: Sparsity-Constrained Wide-Area Control of Power Systems

Wide-area control helps in suppression of inter-area oscillations in electric power systems, but potentially requires a substantial investment into the communication network needed to exchange state information among the power companies [15], [16]. The WAC problem is formulated as an LQR minimization problem. We apply the proposed social optimization and noncooperative game to enable WAC under communication cost constraints. Moreover, we address fair communication cost allocation to the power companies using the Nash Bargaining Solution (NBS) [23].

#### A. Power System Model and Multi-Agent Optimization

The power transmission system with $n$ generators divided into $r$ areas can be modeled by (11)-(12), where each node represents a generator, and an agent (player) represents an area owned by a power company. The disturbance input $w(t)$ is a scalar impulsive disturbance entering the electro-mechanical swing dynamics of any generator, while $D \in \mathbb{R}^{m \times 1}$ is an indicator vector whose entries are all zero except for the one corresponding to the acceleration equation of the generator at which $w(t)$ enters. In this paper, we focus on a traditional 3rd order swing and excitation system model of synchronous generators [16] which employs $m_i = 3$ and $p_i = 1$ for all $i = 1, ..., n$. In the latter case, each generator has 3 states: the generator phase angle (radians) $\delta_i$, the generator rotor velocity (rad/sec) $\omega_i$, and the quadrature-axis internal emf $E_q$. The following [16], it can be shown that the linearized small-signal dynamic model of the networked power system can be expressed in the Kron-reduced form as

$$ \begin{bmatrix} \Delta \delta_m \Delta \omega_m \Delta E_m \end{bmatrix} = \begin{bmatrix} 0 & I & 0 & \Delta \delta \Delta \omega \Delta E \end{bmatrix} \begin{bmatrix} 0 & 0 & \Delta P_m \Delta E_F \end{bmatrix} (15) $$

where $\Delta \delta = \text{col}(\Delta \delta_1, ..., \Delta \delta_n), \Delta \omega = \text{col}(\Delta \omega_1, ..., \Delta \omega_n), \Delta E = \text{col}(\Delta E_1, ..., \Delta E_n), \Delta P_m = \text{col}(\Delta P_{m1}, ..., \Delta P_{mn}), \Delta E_F = \text{col}(\Delta E_{F1}, ..., \Delta E_{FN})$, respectively, represent the...
small-signal changes in phase angle, frequency, excitation voltage, mechanical power input, and excitation voltage input. \( M = \text{diag}(M_1, \ldots, M_n) \) represents the generator inertias, while \( T = \text{diag}(\tau_1, \ldots, \tau_n) \) represents the excitation time constants. The expressions for the various matrices on the RHS can be found in [16]. Equation (\ref{eq:1}) serves as the primary model for wide-area control indicating how the desirable control inputs \( \Delta P_m \) and \( \Delta E_F \) enter the system dynamics. The turbine mechanical power \( \Delta P_m \), however, typically has a much lower bandwidth than needed for oscillation damping. Therefore, for all practical wide-area control designs, \( \Delta P_m \) is treated as zero, and \( \Delta E_F \) is designed via phase, frequency, and voltage feedback available from Phasor Measurement Units (PMUs).

For the sake of simplicity, we assume full state availability for control. If all three states are not measured, then one can employ a phasor state estimator for feedback. The \( A \) matrix for the general case in (\ref{eq:16}) can be obtained by permuting the matrix \( A \) in (\ref{eq:15}) in terms of the sets \( S_1, S_2, \ldots, S_r \) defined in (\ref{eq:7}).

The conventional WAC problem is the centralized LQR optimization that minimizes (\ref{eq:5}), with the solution given by a dense matrix \( K \). We refer to the problem (\ref{eq:6}) applied to WAC as the sparsity-constrained wide-area control (SCWAC). Performance of the proposed social optimization and noncooperative games for SCWAC will be analyzed in Section IV. The cost allocation for SCWAC is described below.

**B. Cost Allocation for SCWAC**

Consider the following cooperative NFG with transferable utility. The players are areas \( \{1, 2, \ldots, r\} \), which cooperate to form links from PMUs to controllers to minimize the energy performance under a global cost constraint. We quantify the allocated cost to each area by the payoff it obtains when it participates in cooperative wide-area control. Several fair payoff allocation methods have been proposed in the literature [24], [25]. In this paper, we employ Nash Bargaining Solution (NBS) due to its computational efficiency in NFGs [23]. The proposed SCWAC cost allocation algorithm proceeds in 3 steps.

1) **Social optimization for SCWAC:**

In the first step, we perform the social optimization (\ref{eq:6}) given the communication cost constraint \( s \), which produces a socially-optimal feedback matrix \( K_{\text{opt}}(s) \). We set \( R_i \) to be the identity matrix so that the energy of every controller has the same weight. The objective function \( J(K) \) in (\ref{eq:6}) is chosen so that all generators arrive at a consensus in their small-signal changes in phase angles and frequencies, as dictated by the physical topology of the network. Since the system is in the Kron-reduced form, the topology of the network is an all-to-all graph. The \( Q \) matrix in (\ref{eq:5}) is obtained from

\[
E_{\text{states}} = \begin{bmatrix} \Delta \delta \\ \Delta \omega \\ \Delta E \end{bmatrix}^T \begin{bmatrix} \hat{L} \\ \hat{L} \\ I \end{bmatrix} \begin{bmatrix} \Delta \delta \\ \Delta \omega \\ \Delta E \end{bmatrix} = x^T (P^T Q' P) x = x^T Q x
\]

where \( P \) is a permutation matrix that rearranges the state vector \( x \) in (\ref{eq:1}) by stacking all the angles first, then all the frequencies, then the excitation voltages, and finally all the remaining states. If more detailed models of synchronous generators are used, then the \( \Delta E^T \Delta E \) term in (\ref{eq:16}) may be replaced by \( (x^-)^T x^- \), where \( x^- \) contains all states except the electro-mechanical states \( \Delta \delta \) and \( \Delta \omega \).

The optimal social energy is defined as

\[
E_{\text{soc}}(s) := J(K_{\text{soc}}(s))
\]

where \( K_{\text{opt}}(s) \) minimizes (\ref{eq:6}) under the cost constraint \( s \), which can be found using either the centralized or the distributed social algorithm in Section II as discussed in Section IV.

2) **The disagreement point:**

The disagreement point (\ref{eq:23}) quantifies the minimum payoff each player expects to obtain when participating in SCWAC. In this case, the selfish objective function for area \( i \) is defined in (\ref{eq:10}), and each area aims to maximize its energy savings when investing in SCWAC given the global communication constraint \( s \). Similarly to the cooperative case, we set \( R_i = I_{N_i \times N_i} \). The matrix \( Q_i \) is chosen to optimize the \( i^{th} \) player’s intra-area energy plus its shared amount of energy due to inter-area oscillations. The intra-area energy of area \( i \) is expressed in the consensus form:

\[
x_i^T Q_i x_i := \sum_{k \in S_i} \sum_{j > k} (\Delta \delta_k - \Delta \delta_j)^2 + \sum_{k \in S_i} \sum_{j > k} (\Delta \omega_k - \Delta \omega_j)^2 + \sum_k \Delta E_k^2
\]

To quantify each area’s contribution to the inter-area energy, we collect the power transfer terms associated with a generator in area \( i \) and a generator in another area, and attribute 1/2 of this energy to area \( i \). Thus, the area \( i \)’s share of the inter-area energy in terms of the phase angle is

\[
\frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\Delta \delta_k - \Delta \delta_j}{n}
\]

The combined intra-area and inter-area energy associated with area \( i \in \{1, \ldots, r\} \) is computed as

\[
x_i^T Q_i x = \Delta \delta_i^T L_i \Delta \delta + \Delta \omega_i^T L_i \Delta \omega + (x^-)^T \mathcal{H}_i x^- \tag{21}
\]

where

\[
L_i = \frac{n - 2n_i}{2} I_n - I_n 1_n 1_n^T (I - I_n) + L I_i \tag{22}
\]

\[
L = \text{diag}(n_1 I_{n_1}, \ldots, n_r I_{n_r}) \tag{23}
\]

\[
I_i = \text{diag}(I_{i,1}, I_{i,2}, \ldots, I_{i,r}) \tag{24}
\]

\[
\mathcal{H} = \text{diag}(H_i^1, H_i^2, \ldots, H_i^r) \tag{25}
\]

\[
I_j = I_{i,j} = I_{n_i \times n_j} \tag{26}
\]

\[
H_j^i = I_{j,i} = I_{(M_j - 2n_j) \times (M_j - 2n_j)} \tag{27}
\]

where \( I_A \) is the indicator function.
To find the disagreement point, we consider two noncooperative games. In both games, the \(i\)th area’s strategic variable is \(K^i\) in (9), and its objective function is \(J_i\) in (10) with \(Q_i\) obtained from (21).

i) Decoupled noncooperative game:

In the first noncooperative game, each area employs only local feedback. Thus for \(i = 1, \ldots, r\),

\[
u_i(t) = -K^{i,\text{diag}}x_i(t) = -\bar{K}^ix_i(t)
\]

where \(\bar{K}^i\) is a block-diagonal matrix with the diagonal blocks given by \(K_{ij}, j \in S_i\), which correspond to local feedback of states within generators in area \(i\) (see (3)), i.e., the feedback is decoupled among the generators. We refer to this game as the decoupled noncooperative game and to the Nash strategies of this game, i.e., \((K^{1*}, \ldots, K^{r*})\) in (13) with \(s = 0\), as the Decoupled Nash Equilibrium (DNE).

DNE is computed by performing the Polishing Step 3 of Algorithm 2 (see [17] for additional details). The energy of area \(i\) at DNE represents the minimum energy that this area can achieve in the absence of cooperation and wide-area feedback while employing its individual objectives. This energy is denoted as

\[
E_{i}^{D} = J_i(\text{DNE})
\]

The total energy of this decentralized noncooperative implementation is

\[
\bar{E}^{D} := \sum_{i=1}^{r} E_{i}^{D}.
\]

ii) Sparsity-constrained noncooperative game:

To determine the energy savings associated with investment into wide-area feedback, we construct a coupled noncooperative game where the generators can send feedback to each other under the same global constraint \(s\) as in Step 1. In the coupled game, the areas invest in WAC while still optimizing their individual objectives. We refer to a NE of this game \((10,21)\) as coupled Nash Equilibrium (CNE). The resulting energy of area \(i\) \((10,21)\) is denoted

\[
E_{i}^{C}(s) := E_i(\text{CNE}; s)
\]

and the global CNE energy is given by

\[
E^{C}(s) := \sum_{i=1}^{r} E_{i}^{C}(s)
\]

iii) Selfish payoffs:

Closed-loop dynamic performance of each area’s power flows is improved by reducing the LQR cost, i.e., by increasing the sparsity-constrained selfish designs, given for area \(i\) by (29) and (31), respectively. This difference determines the energy that area \(i\) saves by utilizing the wide-area feedback in the absence of cooperation under the communication cost constraint \(s\). Thus, we represent the minimum payoff each area expects from WAC by this difference, forming the disagreement point

\[
v_i(s) = E_{i}^{D} - E_{i}^{C}(s), \; i = 1, \ldots, r
\]

However, \(E^{C}(s)\) is not necessarily non-decreasing in \(s\) as demonstrated in Section IV. As a result, while the overall energy of the noncooperative game reduces with \(s\), the disagreement point \((33)\) might have some negative values, and some areas’ payoffs might decrease when \(s\) is relaxed. To satisfy these areas, we propose a disagreement point

\[
v_i(s) = E_{i}^{D} - \min_{s'}\{E_{i}^{C}(s')|s' \leq s\}, \; i = 1, \ldots, r
\]

where \(E_{i}^{C}(s)\) is the best selfish energy performance when the constraint does not exceed \(s\). Note that the disagreement point \((34)\) is hypothetical in a sense that a communication network with the energies \(E_{i}^{C}(s)\) in (34) might not be feasible. However, noncompatible selfish payoffs are often employed in the literature to reflect the player’s subjective preferences and are not required to represent a feasible scenario \((23), (26)\).

3) Payoff and cost allocation:

The energy saved by the social optimization is given by (see (30) and (18))

\[
v_{soc}(s) = \bar{E}^{D} - E_{soc}(s)
\]

which quantifies the total social payoff of cooperative wide-area feedback under the communication cost constraint \(s\). Next, this payoff is divided among the players. Using NBS, the payoff allocated to area \(i\) is given by \((29)\).

\[
\alpha_i = v_i + \frac{v_{soc} - \sum_{i=1}^{r} v_i}{r}
\]

Note that bargaining is successful when

\[
v_{soc}(s) \geq \sum_{i=1}^{r} v_i(s)
\]

In this case, all players receive at least their minimum required payoff \(v_i(s)\). This condition is satisfied for the disagreement point \((33)\) since for each value of \(s\), the energy of the social optimization \((18)\) cannot exceed that of the noncooperative network \((32)\). However, when the disagreement point \((34)\) is used, successful bargaining is not guaranteed for all values of \(s\). Our experiments show that cooperation is successful in practical power system scenarios when the disagreement point \((34)\) is employed. Moreover, when bargaining is achieved for \((34)\), the allocated payoffs are nonnegative for all players.

Finally, the communication costs \(C_i(s), \; i = 1, \ldots, r\) are computed proportionally to the allocated energy payoffs:

\[
C_i(s) = s \cdot \alpha_i / \sum_{i=1}^{r} \alpha_i
\]

Note that \((36), (38)\) theoretically support negative payoffs \(\alpha_i\) and costs \(C_i\), and, thus, payments to some players for cooperating \((23)\). Moreover, due to cooperation, all areas’ payoffs increase by the same amount \(\xi = \alpha_i - v_i\) in \((36)\). If \(\xi\) is small, the social solution does not improve on the selfish solution significantly, i.e., \(E^{C}(s) \approx E_{soc}(s)\). In this case, the communication costs are allocated approximately proportionally to the selfish energy savings \(v_i\) \((17)\). These costs can vary significantly due to areas’ need for feedback as discussed further in section IV.B. On the other hand, when \(\xi\) is large, it dominates \((36)\) (i.e., cooperation significantly improves all areas’ payoffs), and the
communication costs equalize. Finally, in [38], we assume the communication cost constraint \( s \) quantifies the cost of WAC. More generally, this cost can be characterized by the expenditures of wide-area control, a vital component of the overall investment into the communication infrastructure in future power networks [15].

IV. NUMERICAL RESULTS AND PERFORMANCE ANALYSIS

A. Sparsity-Constrained Algorithm Performance for the Australian Power System Model

We validate our results using a 50-bus power system shown in Fig. 1. This power system model consists of 14 synchronous generators, divided into 4 coherent areas, and is a moderately accurate representation of the power grid in south-eastern Australia [27]. The area distribution is shown in different colors in Fig. 1, with the red dots denoting generator buses. Generators 1 to 5 belong to area 1, generators 6 and 7 belong to area 2, generators 8 to 11 belong to area 3, and generators 12 to 14 belong to area 4. The model considers detailed representation of the generator dynamics, with each machine modeled by 14 states. However, since we are primarily interested in the electro-mechanical states, we perform an initial round of model reduction using singular perturbation to eliminate the non-electromechanical states, which have very low participation in the wide-area swing dynamics.

Fig. 2(a) shows the global energy vs. communication cost constraint \( s \) for several centralized and distributed algorithms. Global energies of the centralized optimization [6] using Algorithm 1, the coupled noncooperative game (CNE) [32] (Algorithm 2), the social distributed optimization [14] using Algorithm 2, and of the decoupled game (DNE) [30] are included. We also show the global energy of the iterative dense-feedback method that solves CARE [5] and of the centralized sparsity-promoting ADMM method [6], modified to satisfy the sparsity constraint. Since the energy of the method in [6] is nondecreasing with the sparsity parameter \( \gamma \), a bisection search on \( \gamma \) yields the smallest value of \( \gamma \) for which the off-diagonal cardinality of the feedback matrix produced by the ADMM algorithm satisfies the constraint. The choice of the \( l_1 \)-metric weights for this ADMM implementation is described in [17].

As discussed above, the global energies of all sparsity-constrained methods are theoretically nonincreasing with \( s \). However, due to local convexity of the optimization objective [5], the social algorithms might occasionally converge to a local minimum, thus producing a larger energy as \( s \) increases. If the algorithm results in \( E_{soc}(s_2) > E_{soc}(s_1) \) for \( s_2 > s_1 \), we set \( E_{soc}(s_2) = E_{soc}(s_1) \), which is a suboptimal solution since the feedback matrix for \( s = s_1 \) is a feasible solution for problem [6] when \( s = s_2 \). Note that the energy of the distributed social optimization (using Algorithm 2) closely approximates that of the centralized Algorithm 1 over the \( s \)-range and that the proposed sparsity-constrained methods achieve lower energy than the modified sparsity-promotion ADMM algorithm [6] for most values in \( 4 \leq s \leq 500 \).

We also observe that the global CNE energy tends to those of the DNE and CARE as \( s \) approaches 0 and its largest value \( s = 2223 \), respectively. Finally, the global energies of all social algorithms saturate to the same asymptotic value when \( s \) exceeds 740, implying that the communication cost can be reduced by a factor of 3 relative to the cost of the dense LQR network without compromising the energy performance.

B. Cost Allocation

Fig. 2(b) shows the individual energy objectives of the four areas at CNE [31]. As \( s \) increases, these selfish energies approach those of the dense-feedback game (CARE) [5]. As mentioned in Section III, these individual energies do not necessarily decrease with \( s \). For example, in Fig. 2(b), the DNE energy of area 3, \( E^3_{DNE} \) is slightly smaller than its CNE energy for \( s = 1 \), \( E^3_{CNE}(1) \).

Fig. 3(a) shows the selfish energy savings \( v_i(s) \), \( i = 1, \ldots, 4 \) [33] as well as the social energy savings \( v_{soc}(s) \) [35] computed using Algorithm 2. Note that in this example the disagreement points [33] and [54] result in almost identical numerical results, so we show just the cost allocation for [33]. We observe that bargaining is successful for all \( s \) values since \( \sum_{i=1}^{4} v_i(s) < v_{soc}(s) \), and there is modest payoff increase.
due to cooperation ($\xi = \alpha_i(s) - v_i(s) \leq 1.1$ over the $s$-range). Finally, Fig. (b) shows the allocated communication costs $C_i(s)$ for areas 1 to 4 vs. $s$.

Fig. (a) demonstrates significant disparity in the selfish energies $E_i^s$, the $v_i$ values, and the allocated costs $C_i$ among the areas, which is due to the grid topology. For example, large selfish energy and allocated cost of area 1 can be explained by its large number of generators. However, area size is not the only indicator, e.g., area 4 has fewer generators than area 3, but much larger selfish energy and energy savings (and, thus, the allocated cost), which even exceed those of area 1 in the sparse feedback region. In summary, area 1 and 4 pay a much greater share of the overall network cost than areas 2 and 3 due to greater need for feedback, consistent with relatively steep decline of their selfish energies with $s$.

C. Algorithm Convergence and Implementation Issues

First, we discuss convergence of the proposed games. Since the LQR objective is not necessarily convex, the noncooperative games are not guaranteed to converge to their pure-strategy NE. If the decoupled game converges, the resulting DNE is a local NE due to the local convexity of the LQR objective. On the other hand, the distributed social optimization is an exact potential game, and, thus, a pure NE exists for this game. The optimal solution of the social optimization problem constitutes a NE, although the converse does not necessarily hold due to nonconvexity of the LQR objective.

Second, we focus on the numerical properties of Algorithms 1 and 2. Since the LQR objective does not satisfy the Stable Restricted Hessian condition, convergence of these algorithms is not assured in general. However, if Step 2 of Algorithm 1 converges and yields a stabilizing feedback matrix $\hat{K}$, then Step 3 will also converge due to the local convexity of $J(K)$. At convergence, Step 3 produces a feedback matrix $\hat{K}$ that satisfies the basic feasibility property $\nabla_{\hat{K}}J(\hat{K})|_{\text{supp}(\hat{K})}(\hat{K} = \hat{K}) = 0$, which is a weak necessary condition for the optimality of problem (10). Similar arguments show that Algorithm 2 produces a feedback matrix that satisfies the basic feasibility property of each individual minimization in the CNE problem (13) (or (14) in the case of social optimization). These observations demonstrate that convergence properties of proposed algorithms resemble those of the ADMM-based methods. In both cases, while theoretical guarantees are not always feasible, extensive numerical experience demonstrates that the algorithms converge and provide desirable minimization solutions over a range of sparsity parameters. For example, the proposed sparsity-constrained algorithms also converge and exhibit similar performance and complexity trends to those shown in this paper for the New England power system model used in [6].

Next, we describe the algorithm implementation details for the results shown in Fig. (a). We found that the proposed algorithms can converge to different stabilizing feedback matrices given different initial settings, and the energies of these solutions can differ significantly for the Algorithms 1 and 2, but are very similar to each other for the Decoupled game (28). We have employed $\epsilon_{\text{abs}} = \epsilon_{\text{rel}} = 10^{-4}$ for Algorithm 1 and 2, $\epsilon_1 = 10^{-4}$ for Algorithm 1 and $\epsilon_2 = 10^{-3}$ for Algorithm 2 to achieve comparable performance for distributed and centralized social optimizations. Both algorithms were initialized for $s = 1$ with a stabilizing matrix $\hat{K}_0$ obtained by preserving the block-diagonal entries of the standard LQR feedback matrix, and setting other entries to zero. We found that this initialization produced the lowest energies over the entire $s$-range. If a stable block-diagonal matrix cannot be found, $\hat{K}_0$ can be obtained using the sparsity-promotion algorithm [21] with the largest $\gamma$ that produces a stabilizing feedback matrix. (Similarly, stabilizing feedback might not exist for small values of the constraint $s$.) In addition, Algorithm 2 has the best performance when the initial link settings $s_i$ are chosen proportionally to the number of nodes $n_i$ in (7).

Finally, we found that the computational load of the polishing step dominates the overall runtime for both algorithms, and the Newton step using the CG method is the most computation-intensive operation, which has polynomial complexity in $s$ and the number of states. In our experiments, all algorithms in Fig. (a) converged in less than $10^3$ seconds for any value of $s$, although this did not include the bisection search time for the modified ADMM method, which is very computation-intensive. Moreover, the distributed social implementation using Algorithm 2 converged much faster than the centralized method (Algorithm 1).

V. Conclusion

Sparsity-constrained LQR optimization with linear state feedback was investigated. A centralized social optimization...
tion method based on the GraSP algorithm and restricted
Newton step was proposed. Moreover, these methods were
combined with iterative gradient descent to develop a sparsity-
constrained noncooperative linear-quadratic game with dis-
tributed computation. By converting this game into a potential
game with a common optimization objective for all players, we
have developed a sparsity-constrained distributed social opti-
mization algorithm. Performance and convergence properties
of proposed algorithms were demonstrated for a 50-bus power
system model divided into 4 areas.

REFERENCES
[1] J. Sztipanovits, X. Koutsoukos, G. Karsai, N. Kottenstette, P. Antsaklis,
V. Gupta, B. Goodwine, J. Baras, and S. Wang, “Toward a science of
cyber–physical system integration,” Proceedings of the IEEE, vol. 100,
no. 1, pp. 29–44, 2012.
[2] T. Başar and G. J. Olsder, Dynamic noncooperative game theory. SIAM,
1995, vol. 200.
[3] F. L. Lewis and V. L. Syrmos, Optimal control. John Wiley & Sons,
1995.
[4] D. Dukh and D. Russell, “A global theory for linear-quadratic differen-
tial games,” Journal of Mathematical Analysis and Applications, vol. 33,
no. 1, pp. 96–123, 1971.
[5] H. Mukaidani, “A numerical analysis of the Nash strategy for weakly
coupled large-scale systems,” IEEE Transactions on Automatic Control,
vol. 51, no. 8, pp. 1371–1377, Aug. 2006.
[6] F. Dörfler, M. R. Jovanović, M. Chertkov, and F. Bullo, “Sparse and
optimal wide-area damping control in power networks,” in American
Control Conference (ACC), 2013, pp. 4289–4294.
[7] M. Fardad, F. Lin, and M. R. Jovanović, “On the optimal design of
structured feedback gains for interconnected systems,” in Decision
and Control, 2009 held jointly with the 2009 28th Chinese Control
Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference
on, IEEE, 2009, pp. 978–983.
[8] F. Lin, M. Fardad, and M. R. Jovanović, “Augmented Lagrangian
approach to design of structured optimal state feedback gains,” IEEE
Transactions on Automatic Control, vol. 56, no. 12, pp. 2923–2929,
2011.
[9] S. Bahmani, B. Raj, and P. T. Boufounos, “Greedy sparsity-constrained
optimization,” The Journal of Machine Learning Research, vol. 14, no. 1,
pp. 807–841, 2013.
[10] A. Beck and Y. C. Eldar, “Sparsity constrained nonlinear optimization:
Optimality conditions and algorithms,” SIAM Journal on Optimization,
vol. 23, no. 3, pp. 1480–1509, 2013.
[11] A. Lamperski and L. Lessard, “Optimal decentralized state-feedback
control with sparsity and delays,” Automatica, vol. 58, pp. 143–151,
2015.
[12] J. F. Mota, J. M. Xavier, P. M. Aguia, and M. Puschel, “D-ADMM:
A communication-efficient distributed algorithm for separable optimiza-
tion,” IEEE Transactions on Signal Processing, vol. 61, no. 10, pp.
2718–2723, 2013.
[13] J. J. Ralliff, S. A. Burden, and S. S. Sastry, “Characterization and com-
cputation of local Nash equilibria in continuous games,” in 51st An-
nual Allerton Conference on Communication, Control, and Computing
(Allerton). IEEE, 2013, pp. 917–924.
[14] N. Li and J. R. Marden, “Designing games for distributed optimization,”
IEEE Journal of Selected Topics in Signal Processing, vol. 7, no. 2, pp.
230–242, 2013.
[15] P. T. Myrda and K. Koellner, “NASPnet-the Internet for synchrophasors,” in IEEE 43rd Hawaii International Conference on System Sciences
(HICSS), 2010, pp. 1–6.
[16] A. Chakrabortty and P. Khargonekar, “Introduction to wide-area moni-
toring and control,” in American Control Conference, DC, 2013.
[17] F. Lian, A. Duel-Hallen, and A. Chakrabortty, “Ensuring economic
fairness in wide-area control for power systems via game theory,” in
American Control Conference, 2016.
[18] M. F. Beeler, D. Simchi-Levi, and C. Barnhart, “Network cost allocation
games on growing electricity grids,” IIE Annual Conference Proceed-
ing, vol. 3238, 2014.
[19] F. Lian, A. Duel-Hallen, and A. Chakrabortty, “Cost allocation strategies
for wide-area control of power systems using Nash bargaining solution,” in
IEEE 53rd Annual Conference on Decision and Control, Dec 2014, pp.
1701–1706.