Consensus Control for Linear Systems with Optimal Energy Cost

Han Zhang\textsuperscript{a}, Xiaoming Hu\textsuperscript{a},

\textsuperscript{a}Department of Mathematics, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden

Abstract

In this paper, we design an optimal energy cost controller for linear systems asymptotic consensus given the topology of the graph. The controller depends only on relative information of the agents. Since finding the control gain for such controller is hard, we focus on finding an optimal controller among a classical family of controllers which is based on Algebraic Riccati Equation (ARE) and guarantees asymptotic consensus. Through analysis, we find that the energy cost is bounded by an interval and hence we minimize the upper bound. In order to do that, there are two classes of variables that need to be optimized: the control gain and the edge weights of the graph and are hence designed from two perspectives. A suboptimal control gain is obtained by choosing \( Q = 0 \) in the ARE. Negative edge weights are allowed, and the problem is formulated as a Semi-definite Programming (SDP) problem. Having negative edge weights means that "competitions" between the agents are allowed. The motivation behind this setting is to have a better system performance. We provide a different proof compared to Thunberg & Hu (2016) from the angle of optimization and show that the lowest control energy cost is reached when the graph is complete and with equal edge weights. Furthermore, two sufficient conditions for the existence of negative optimal edge weights realization are given. In addition, we provide a distributed way of solving the SDP problem when the graph topology is regular.

Key words: Consensus control, multi-agent systems, optimal control, semi-definite programming; distributed optimization

1 Introduction

Cooperative control for multi-agent systems has been intensively studied in the past decade, with wide applications on smart power grid, robots and so on. The cooperative control methods usually imply a distributed manner, which make the systems enjoy many advantages such as being robust and economic.

Consensus is an important topic in the research of cooperative control for multi-agent. The goal is to let the states or the outputs of all agents become the same by control laws that depend on the information of the agent and its neighbours. In this paper, we consider the case of asymptotic consensus for linear systems. The goal of this paper is to design distributed controllers using only the relative information between the agents and with minimal control energy cost such that all the system states will eventually become the same as time goes infinity.

In this paper, we focus on designing the optimal energy cost controller for the linear systems so that they can reach asymptotic consensus. It is well-known that the asymptotic consensus for linear systems is equivalent to regulating \( N - 1 \) systems with the dynamics \( \dot{x}_i = Ax_i + \lambda_i B u_i, 2 \leq i \leq N, \) where \( \lambda_i \) is the \( i \)th smallest eigenvalue of the Laplacian matrix, see (Fax & Murray, 2004), (Zhang et al., 2011). The area is well-studied, for example, in (Borrelli & Keviczky, 2008), the authors design linear quadratic regulators for identical linear systems when the graph topology is given. Though it is a quite similar to what we undertake here, but as mentioned above, the equivalent problem is to regulate \( N - 1 \) different systems with minimum energy, and hence is not the same. The authors of (Rogge et al., 2010) consider the problem as a quadratic optimal control problem on a ring network while we consider the case of a graph with arbitrary topologies. Augmented Lagrangian approach is used in (Lin et al., 2011) to design a structured distributed controller so that the \( H_2 \) norm of the noisy systems is minimized. Deshpande et al. (2011) consider a similar problem and use a two-step approach to design the control law. But their controller does not only use relative information of the agents, but also uses the agents’
own states. Cao & Ren (2010) study the optimal consensus of the single-integrators for both discrete-time and continuous-time case. In (Lin et al., 2013), alternating direction method of multipliers (ADMM) is used to minimize the $H_2$ norm so that an optimal sparse feedback gain is obtained. In (Thunberg & Hu, 2016), a “topology free” control energy minimization problem is considered and the distributed energy-optimal control corresponds to a complete graph with equal edge weights.

On the other hand, distributed optimization has attracted great attentions these years due to the wide applications in the network. Compared to the abundant results on distributed optimization in real vector spaces, for instances, (Yi et al., 2015), (Annergren et al., 2014), (Lou et al., 2014), (Nedic & Ozdaglar, 2009), (Nedic et al., 2010), the results on distributed SDP still remain limited. Dall Anese et al. (2013) considered an optimal power flow problem in power grids and relax it into an SDP problem. Then ADMM is used to solve the problem in a distributed manner. In (Simonetto & Leus, 2014), a sensor localization problem is considered. Relaxation towards an SDP and ADMM is used to solve the problem distributedly as well. Pakazad et al. (2014) analyse the robustness of interconnected uncertain systems and linear matrix inequalities (LMIs) are reformulated in SDP. Chordal sparsity structure is assumed among the data matrices of the SDP and hence the problem is decomposed and solved distributedly using proximal splitting method. In (Pakazad et al., 2015), coupled SDPs with tree structures are considered. A distributed primal-dual interior point method is proposed to solve the coupled SDPs. The aforementioned work all utilize the idea of “decomposition” somehow but in this paper we treat the SDP in a different manner: reaching an optimal consensus in the intersection of convex feasible sets. Also, what makes our work different from existing distributed optimization problem is that our problem motivates from optimizing a parameter of a graph and the communication network of the distributed optimization algorithm is actually the physical network itself, while most distributed optimization algorithms relax and decompose the original problem and “design” the communication network according to the structure of the decomposed problem, see Pakazad et al. (2015) as an example.

The main contribution of this paper is the construction of an optimal energy controller that depends only on the relative information between the agents. The controller has two classes of variables that need to be determined: the control gain and the edge weights of the graph. Similar to (Borrelli & Keviczky, 2008), computing the optimal control gain for the controller is hard, thus we focus on finding an optimal controller among a classical family of controller designs based on ARE and guarantees asymptotic consensus. Through analysis, we found that the energy cost is bounded by an interval and hence we minimize the upper bound. A suboptimal control gain is obtained by choosing $Q = 0$ in the ARE; the edge weights of the graph is optimized by solving an underlying SDP. The controller that we designed enjoy several favourable properties:

(1) The controller coincides with the optimal control in (Thunberg & Hu, 2016) when the graph is complete. It has been pointed out in (Thunberg & Hu, 2016) that any other distributed control laws constructed by Laplacian matrices that do not correspond to complete graphs with equal edge weights are suboptimal.
(2) When optimizing the edge weights, “competitions” are allowed between the connected agents. By doing so, the feasible region of the optimization problem is enlarged, and hence a smaller control energy cost might be obtained. We offer two sufficient conditions for when will “competitions” happen between agents. These two conditions help to determine whether the two agents will compete if we add a connection between them based on the old optimal solution.
(3) When the graph topology is regular, namely, every node has the same number of neighbours, the controller can be calculated in a distributed manner.

The rest of this paper is organized as follows. In Section 2, some preliminaries and notations are presented. In Section 3 we introduce the problem formulation. The design of the control gain is presented in Section 4.1 and 4.2. In Section 4.3, the edge weights design is formulated as an SDP problem and we provide a different proof compared to (Thunberg & Hu, 2016) from the angle of optimization and show that the lowest control energy cost is reached when the graph is complete and with equal edge weights. Two sufficient conditions on the existence of the negative optimal edge weight are presented. In Section 4.4, a distributed way of computing the optimal edge weights is presented when the graph is regular. Finally, we conclude the paper and describe some future work in Section 5.

2 Notations and Preliminaries

In this section, we provide a brief introduction about some notations that we are going to use in the rest of the paper, as well as some basic knowledge about SDP.

We denote $\mathbf{1}$ as an $N$ dimensional all-one column vector. $\mathbf{0}$ is denoted as an N dimensional all-zero matrix. The element on the $i$th row and $j$th column of any matrix $D$ is expressed as $[D]_{ij}$. $D_1 \succeq D_2$ and $G_1 \preceq G_2$ mean that $D_1 - D_2$ and $G_2 - G_1$ are positive semi-definite. $\otimes$ denotes the Kronecker product. $\| \cdot \|$ denotes 2-norm of matrice or vectors. $\| \cdot \|_F$ denotes the Frobenius norm of a matrix. We use $\cdot$ to denote the number of the elements of a set. And any notation with the superscript * denotes the optimal solution to the corresponding optimization problem. $tr(\cdot)$ denotes the trace of a matrix. If
\(D, G \in S_n^+\) are positive definite matrices, then \(\text{tr}(DG)\) is the inner-product between \(D\) and \(G\). It also holds that \(\nabla_{\text{tr}}(DG) = G\).

An edge-weighted undirected graph \(\mathcal{G}(V, \mathcal{E}, \mathcal{W})\) is composed of a node set \(V = \{1, 2, \cdots, N\}\), an edge set \((i, j) \in \mathcal{E}, i, j \in V\) which describes the connection topology between the nodes and the edge weight set \(w_{ij} \in \mathcal{W}, i, j \in V\) which includes all the weights of the corresponding edges. To abbreviate the notation, we label the edges with numbers. For example, an edge with label \(l\) is denoted as \(l \in \mathcal{E}\). On the other hand, seen from the nodes’ perspective, the set of edges that is connected to node \(i\) is denoted as \(\mathcal{E}(i)\), which can be interpreted as communication channels of node \(i\). Note that \(\mathcal{E} = \bigcup_{i \in V} \mathcal{E}(i)\). Similarly, the set of the edge weights belong to node \(i\) is denoted as \(\mathcal{W}(i)\). \(\mathcal{N}(i)\) denotes the neighbour vertices set of node \(i\).

Note that in this paper, we consider undirected graphs, hence the edge-weighted Laplacian matrix \(L_w\) is symmetric and defined as

\[
[L_w]_{ij} = \begin{cases} 
\sum_l w_{il} & \text{if } i = j \text{ and } (i, l) \in \mathcal{E} \\
-w_{ij} & \text{if } i \neq j \text{ and } (i, j) \in \mathcal{E} \\
0 & \text{otherwise}
\end{cases}
\]

\(\Leftrightarrow L_w = \sum_{k \in \mathcal{E}} w_k E_k\),

where \(k\) is the label of the edges, \(w_k \in \mathcal{W}, \forall k \in \mathcal{E}\) are the edge weights. If node \(i\) and \(j\) are connected via edge \(k\), then

\[
[E_k]_{ii} = [E_k]_{jj} = 1, \quad [E_k]_{ij} = [E_k]_{ji} = -1,
\]

and the other elements of \(E_k\) are zero. For a connected graph, the eigenvalues of \(L_w\) is denoted as 0 = \(\lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N\).

For any symmetric matrix \(G \in S^n\), \(\text{svec}(G)\) is defined as

\[
\text{svec}(G) = [G_{11}, \sqrt{2} G_{21}, \cdots, \sqrt{2} G_{n1}, G_{22}, \sqrt{2} G_{32}, \cdots, \sqrt{2} G_{n2}, \cdots, G_{nn}]^T.
\]

It follows from the above definition that

\[
\text{tr}(DG) = \text{svec}(D)^T \text{svec}(G), \quad \forall D, G \in S^n.
\]

The **Symmetric Kronecker Product** between two matrices \(G\) and \(D\) is defined by the following identity

\[
(R_1 \otimes R_2)\text{svec}(G) = \frac{1}{2} \text{svec}(R_2G R_1^T + R_1GR_2^T),
\]

where \(G \in S^n\), but \(R_1\) and \(R_2\) is not necessarily symmetric. For more details about \(\text{svec}()\) and Symmetric Kronecker Product can be found in Schäcke (2013) and Alizadeh et al. (1998).

Now we introduce some basics about SDP. An SDP problem in standard form can be expressed as

\[
\begin{align*}
\text{maximize} & \quad \sum_{i=1}^{m} c_i \xi_i \\
\text{subject to} & \quad \sum_{i=1}^{m} \xi_i G_i \succeq R.
\end{align*}
\]

Introducing the Lagrangian multiplier \(\Phi \succeq 0\), the Lagrangian function of (1) can be written as

\[
\mathcal{L}(\xi, \Phi) = \sum_{i=1}^{m} c_i \xi_i + \text{tr}(\Phi(\sum_{i=1}^{m} \xi_i G_i - R))
\]

Taking the gradient of the Lagrangian function, we get its dual problem

\[
\begin{align*}
\text{minimize} & \quad \text{tr}(R \Phi) \\
\text{subject to} & \quad \text{tr}(G_i \Phi) = c_i, \quad i = 1, \cdots, m \\
& \quad \Phi \succeq 0.
\end{align*}
\]

### 3 Problem Formulation

In this paper, we consider the following linear multi-agent system, where each agent has the following dynamics:

\[
\dot{x}_i = Ax_i + Bu_i, \quad i = 1, \cdots, N,
\]

where \(x_i \in \mathbb{R}^n, \forall i\) and they are connected through an edge-weighted graph \(\mathcal{G}(V, \mathcal{E}, \mathcal{W})\).

Suppose that \((A, B)\) is stabilizable, \(A\) has no eigenvalue on the imaginary axis. We want each \(x_i\) to reach consensus asymptotically, namely, \(\|x_i(t) - x_j(t)\| \to 0, \forall i, j\) as \(t \to \infty\), while minimizing the control energy cost. Plus, the control should be fully distributed and the control should only depend on relative information, i.e., the control must have the form

\[
u_i = K \sum_{j \in \mathcal{N}(i)} w_{ij} (x_i - x_j).
\]

Note here that, compared to the classical definition, \(w_{ij}\) does not have to be positive, meaning that we allow some of the agents to compete against each other. The reason for us to allow the edge weights to be negative will be described later.
Now we formulate the problem as an infinite horizon Linear Quadratic (LQ) optimal control problem:

\[
\begin{align*}
\text{minimize} & \quad J(K, \{w_{ij}\}) = \int_0^\infty U^T U \, dt \\
\text{subject to} & \quad \dot{X} = (I_N \otimes A)X + (I_N \otimes B)U, \quad (LQ1) \\
& \quad U = (L_w \otimes K)X, \\
& \quad \|x_i(t) - x_j(t)\| \to 0, \ t \to \infty, \ \forall i,j,
\end{align*}
\]

where \( X = [x_1^T, \ldots, x_N^T]^T, U = [u_1^T, \ldots, u_N^T]^T, \) and \( L_w \) is the weighted Laplacian matrix of the edge-weighted graph \( \mathcal{G}(V, E, W) \).

The problem \((LQ1)\) is well-posed though \( \lim_{t \to \infty} x_i(t) = \infty \) when \( A \) is not stable, namely, \( J(K^*, \{w_{ij}\}) < \infty \). To see this, we simplify the system in \((LQ1)\) by applying a coordinate change \( \tilde{X} = (T^T \otimes I_n)X, \) where \( T^T T = I \)

\[
\tilde{X} = (I_N \otimes A)\tilde{X} + (I_N \otimes B)U,
\]

and therefore \( \lim_{t \to \infty} \tilde{x}_i(t) = 0, i \geq 2 \) if consensus is reached in the original system \((2)\). On the other hand, if \( \lim_{t \to \infty} \tilde{x}_i(t) = 0, i \geq 2, \) then \( \lim_{t \to \infty} X(t) = \lim_{t \to \infty} (T \otimes I_n)\tilde{X}(t) = \lim_{t \to \infty} (\frac{1}{\sqrt{N}}1, t_2, \ldots, t_N) \otimes I_n)([1, 0, \ldots, 0]^T \otimes \tilde{x}_1(t)) = \frac{1}{\sqrt{N}}1 \otimes \tilde{x}_1(t). \) Therefore, letting \((2)\) reach asymptotic consensus is equivalent to regulating \( \tilde{x}_i, i = 2, \ldots, N. \) Furthermore, the cost function \( J \) in \((LQ1)\) reads:

\[
J(K, \{w_{ij}\}) = \int_0^\infty U^T U \, dt = \int_0^\infty X^T (I_N^2 \otimes K^T K)X \, dt
\]

\[
= \int_0^\infty \tilde{X}^T (T^T \otimes I_n)(I_N^2 \otimes K^T K)(T \otimes I_n)\tilde{X} \, dt
\]

\[
= \int_0^\infty \tilde{X}^T (T^T L_w T \otimes I_n)(L_w \otimes K^T K)\tilde{X} \, dt
\]

\[
= \int_0^\infty \tilde{X}^T (\tilde{A}_w^2 \otimes K^T K)\tilde{X} \, dt,
\]

(4)

Note that \( \Lambda_L = \text{diag}(0, \lambda_2, \ldots, \lambda_N) \), hence the \( \tilde{x}_1 \) does not appear in the cost function \( J \) and we can remove everything about \( \tilde{x}_1 \) and simplify the problem \((LQ1)\) as:

\[
\begin{align*}
\text{minimize} & \quad J(K, \{w_{ij}\}) = \int_0^\infty \tilde{U}^T \tilde{U} \, dt \\
\text{subject to} & \quad \dot{\tilde{X}} = (I_N \otimes A)\tilde{X} + (I_N \otimes B)\tilde{U}, \\
& \quad \tilde{U} = (\tilde{A}_w^2 \otimes K)\tilde{X},
\end{align*}
\]

(5)

(6)
By rewriting (6), we get
\[ A^T P + PA - PPB^T B^T P = -Q, \]
\[ \Leftrightarrow (A^T P + PA^i_d) - Q = -(2\sigma_i - 1)PPB^T B^T P. \] 
(7)

Suppose \((\lambda_d, v_d)\) is an eigen-pair of \(A^i_d\); then multiply \(v^H_d\) on the left and \(v_d\) on the right on both hand sides of (7), it follows that

\[ 2Re(\lambda_d) \cdot v^H_d P v_d = -v^H_d Q v_d - (2\sigma_i - 1)v^H_d PPB^T B^T P v_d. \]

where \(v^H_d\) denotes the conjugate transpose of \(v_d\). The eigen-pair \((\lambda_d, v_d)\) must satisfy one of the following three cases:

1. If \(v^H_d P v_d = 0\), then \(P v_d = 0\). Hence \(A^i_d v_d = (A - \sigma_i B B^T P)v_d = A v_d = \lambda_d v_d\). On the other hand, \(A^i_d (v_d) = (A - \sigma_i B B^T P)v_d = A v_d = \lambda_d v_d\), thus \((A_d, v_d)\) is also an eigen-pair of \(A_p\). We know that \(A_p\) is Hurwitz, and hence \(Re(\lambda_d) < 0\).

2. If \(v^H_d P v_d > 0\) and right hand side of (8) is strictly negative, then it is easy to conclude that \(Re(\lambda_d) < 0\).

3. If \(v^H_d P v_d > 0\) and right hand side of (8) is zero, then it follows that \(Re(\lambda_d) = 0\). It also follows that \(v^H_d Q v_d = 0\) and \(v^H_d P B B^T B^T P v_d = 0\) since both \(Q\) and \(P B B^T P\) is positive semi-definite. Hence \(B B^T P v_d = 0\). Then we can also conclude that \(A^i_d v_d = A v_d = \lambda_d v_d\) and hence \((\lambda_d, v_d)\) is also an eigen-pair of \(A_p\). We know that \(A_p\) is Hurwitz, thus \(Re(\lambda_d) < 0\). This contradicts \(Re(\lambda_d) = 0\) derived from (8). Thus this case will never happen.

Since any eigenvalue \(\lambda_d\) have \(Re(\lambda_d) < 0\), then \(A^i_d\) is Hurwitz. Therefore control (5) will let the multi-agent system (2) reach consensus asymptotically. \(\square\)

Note that the control gain (5) depends on the choice of \(Q\), hence we would like to choose a good \(Q\) so that the control gain \(K\) should optimize the energy cost.

4.2 Interval Bound on the Energy Cost

Plugging the dynamics of the closed-loop system and the control gain (5) into the cost function \(J(K, \{w_{ij}\})\), we get

\[ J(K, \{w_{ij}\}) = \int_0^\infty \dot{X}^T \left( A^2_L \otimes K^T K \right) \dot{X} \, dt \]

\[ = \int_0^\infty \dot{X}^T \left( A^2_L \otimes \frac{1}{\lambda_2^2} PPB^T B^T P \right) \dot{X} \, dt = \dot{X}^T_0 \left( \frac{1}{\lambda_2^2} PPB^T B^T P \right) \dot{X}_0, \]

where \(\dot{X}_0 = \int_0^\infty e^{(I_{N-1} \otimes A - \Sigma^L \otimes B B^T P)T} (\Sigma^L \otimes PPB^T B^T P) \, dt\), \(\Sigma^L = diag(\sigma_2, \cdots, \sigma_N)\) and \(\sigma_i = \lambda_i/\lambda_2\). Note that \(H\) is block diagonal, i.e., \(H = diag(H_2, \cdots, H_N)\), where

\[ H_i = \int_0^\infty e^{A^i_d t} \sigma_i^2 P B B^T P e^{A^i_d t} \, dt. \]

**Theorem 2** \(P_0 \leq H_i \leq \frac{\sigma_i^2}{2\sigma_i - 1} P\), where \(P_0\) is the solution to the ARE (6) when \(Q = 0\) and \(P\) is the solution to the ARE (6) for any \(Q \geq 0\).

**PROOF.** We first show the upper bound. Note that \(H_i\) has the form of (10), hence it is actually the analytic solution to the following Lyapunov equation

\[ A^i_d H_i + H_i A^i_d = -\sigma_i^2 P B B^T P. \]

(11)

Multiply \(\frac{\sigma_i^2}{2\sigma_i - 1}\) on both hand sides of (7) and subtract (11), we have

\[ A^i_d (\frac{\sigma_i^2}{2\sigma_i - 1} P - H_i) + (\frac{\sigma_i^2}{2\sigma_i - 1} P - H_i) A^i_d = -\frac{\sigma_i^2}{2\sigma_i - 1} Q, \]

which is a Lyapunov equation. Consequently, since \(Q \geq 0\) and \(A^i_d\) is stable, it must hold that

\[ \frac{\sigma_i^2}{2\sigma_i - 1} P - H_i = \int_0^\infty e^{A^i_d t} \frac{\sigma_i^2}{2\sigma_i - 1} Q e^{A^i_d t} \, dt \geq 0. \]

(12)

Hence \(\frac{\sigma_i^2}{2\sigma_i - 1} P \geq H_i\).

Now we show the lower bound. When \(Q = 0\), the solution \(P_0\) must satisfy

\[ A^T P_0 + P_0 A - P_0 B B^T B^T P_0 = 0 \]

\[ \Leftrightarrow A^i_d T P_0 + P_0 A^i_d = -\sigma_i P B B^T P_0 - \sigma_i P_0 B B^T P \]

\[ + P_0 B B^T P_0. \]

(13)

Subtract (11) by (13), we get

\[ A^i_d T (H_i - P_0) + (H_i - P_0) A^i_d \]

\[ = -(\sigma_i P - P_0) B B^T (\sigma_i P - P_0). \]

Also, since \((\sigma_i P - P_0) B B^T (\sigma_i P - P_0) \geq 0\) and \(A^i_d\) is Hurwitz, it must hold that

\[ H_i - P_0 = \int_0^\infty e^{A^i_d t} (\sigma_i P - P_0) B B^T (\sigma_i P - P_0) e^{A^i_d t} \, dt \]

\[ \geq 0. \]
Hence $H_i \succeq P_0$. \hfill \Box

Note that the interval bound will degenerate to a "point" if $Q = 0$ is chosen and when the graph is complete, namely, in this case $P = P_0$ and $\sigma_i = 1$, $i = 2, \cdots, N$. Since by definition $\sigma_i \geq 1$, then $\frac{\sigma_i^2}{\sigma_i - 1}$ is monotonously increasing with respect to $\sigma_i$, hence $P_0 \preceq H_i \preceq \frac{\sigma_i^2}{\sigma_i - 1} P \preceq \frac{\sigma_N^2}{\sigma_N - 1} P$.

Because the optimal control energy cost has a lower bound, the distance between the optimal control energy cost and the upper bound is also bounded and hence it is a rational choice to optimize the upper bound, i.e., $\frac{\sigma_N^2}{\sigma_N - 1} P$, since it is hard to optimize the real energy cost.

Note that the upper bound consists of two variables: $\sigma_N$ and $P$ and hence is optimized from two perspectives. From the perspective of $P$, based on the monotonicity of the solution to the ARE [Willems (1971)], (namely, if $Q_1 \succeq Q_2$, then the corresponding solution holds $P_1 \succeq P_2$), the upper bound is minimized by choosing $Q = 0$. More precisely, when $Q = 0$, the right hand side of (12) becomes zero and hence $H_i = \frac{\sigma_i^2}{\sigma_i - 1} P_0$.

Now that we have optimized the upper bound from the perspective of $P$, what remains to optimize is $\sigma_N$, which is determined by the edge weights. This will be done in the next section.

4.3 Optimizing the Edge Weights

$\frac{\sigma_N^2}{\sigma_N - 1}$ is monotonously increasing since $\sigma_N \geq 1$, hence minimizing $\frac{\sigma_N^2}{\sigma_N - 1}$ is equivalent to minimizing $\sigma_N$.

Remark 3 $\sigma_N = \frac{\lambda_N}{\lambda_2}$ is so-called “synchronizability” in the network synchronization field on physics, and its optimization has been widely studied in the past. Jalili (2013) gives a thorough survey on this topic. Though there has been a lot of approaches to enhance synchronizability in the field of physics [Jalili, 2013], heuristic approaches are used to adjust the edge weights and the effectiveness of the approaches are shown empirically by numerical examples. We would like to know what is the exact optimal edge weight realization to minimize the synchronizability. Since the problem involves the eigenvalues of the matrix $L_w$, it is natural to formulate the problem as an SDP problem. \hfill \Box

Recall that in the problem formulation, $L_w$ is symmetric and negative edge weights $w_{ij}$ are allowed. Furthermore, it is assumed that there is a finite communication resource in the network, namely, $\sum_{k \in E} w_{ik}^2 = \text{const}$. Without loosing generality, we assume the constant to be 1. Of course, the optimized edge-weighted graph $G$ should be connected, namely, the second smallest eigenvalue of the Laplacian matrix $L_w$ should be strictly positive. To ensure the existence of feasible edge weight realization, we make the assumption that the unweighted graph represented by the unweighted Laplacian matrix $L_u = \sum_{k \in E} E_k$ is connected. Thus, the problem can be formulated as the following optimization problem:

\[
\begin{align*}
\text{minimize} & \quad \frac{\lambda_N}{\lambda_2} \\
\text{subject to} & \quad \lambda_2 I - \mu 1 1^T \preceq \sum_{k \in E} w_k E_k \preceq \lambda_N I, \\
& \quad \sum_{k=1}^{\lvert E \rvert} w_k^2 = 1, \\
& \quad \lambda_2 > 0.
\end{align*}
\]

(P1)

In (P1), the variable $\mu$ is used to shift the zero eigenvalue of the Laplacian matrix $L_w$ with its eigenvector $1$. When the optimal value is reached, $\lambda_2^*$ would be the smallest eigenvalue of $\sum_{k=1}^{\lvert E \rvert} w_k E_k + \mu 1 1^T$. By using the property that for any semi-positive definite matrix $G$, we have $\beta_1 I \preceq G \preceq \beta_n I$, where $\beta_1, \beta_n$ is the smallest and biggest eigenvalue of $G$ respectively, we get the above constraints.

Remark 4 The reason we put a free variable $\mu$ in the constraint $\lambda_2 I - \mu 1 1^T \preceq \sum_{k=1}^{\lvert E \rvert} w_k E_k$ instead of writing the constraint as $\lambda_2 (I - \frac{1}{\lambda_2} 1 1^T) \preceq \sum_{k=1}^{\lvert E \rvert} w_k E_k$ is that the latter formulation does not have any interior point in the feasible domain, i.e., every feasible point lies on the boundary. Later we will consider the dual problem and we want to use the Slater’s condition to show the strong duality holds, which requires the primal problem to have at least one strictly feasible solution. Note that when optimal value is reached, $\mu$ does not have to be $\lambda_2^*/N$. In fact, given an optimal edge weight realization and $\lambda_2^*$, any $\mu \geq \lambda_2^*/N$ is an optimal solution.

Remark 5 Note that in (P1), we allow the agents to "compete" by not introducing the constraints $w_{ik} \geq 0, k = 1, \cdots, \lvert E \rvert$ into the problem. If we added the constraints of non-negative edge weights, the feasible domain would shrink and hence potentially gives a larger optimal value. Thus, allowing the agents to compete will give an at least as good, sometimes better system performance, which is quite contrast to intuition. \hfill \Box

Though the problem (P1) is not convex, it can be rewritten into a convex problem. By applying variable change $t = \lambda_N/\lambda_2$, $\gamma = 1/\lambda_2$, $\mu/\lambda_2 = y_0$, $w_k/\lambda_2 = y_k$ and $1/\lambda_2 = \gamma$, we change the problem into a semidefinite
programming (SDP) problem:

\[
\begin{align*}
\text{minimize} & \quad t \\
\text{subject to} & \quad I - y_0 11^T \preceq \sum_{k=1}^{\vert \mathcal{E} \vert} y_k E_k \preceq t I, \quad (P2) \\
& \quad \gamma > 0.
\end{align*}
\]

Note that the last constraint \(\gamma > 0\) in (P2) is redundant since it does not affect the other constraints and the objective function. We can just remove the constraint and get a classic SDP problem. It is also worth noticing that with the aforementioned change of variables, the control (3) converges to the following form:

\[
u_i = -B^TP_0 \sum_{j \in N(i)} y_{ij} (x_i - x_j).
\]

This means that \(J\left(-\frac{1}{\lambda} B^TP_0, \{w_{ij}\}\right) = J(-B^TP_0, \{y_{ij}\})\). (It depends on how we denote \(K\!).

We know that the optimal value of the primal problem can always be attained. Hence if we are able to show that the Slater’s condition holds for (P2), i.e., there exists a solution \((t, y_k, y_0)\) which is strictly feasible, then the strong duality holds. Note that for the constraint \(\sum_{k=1}^{\vert \mathcal{E} \vert} y_k E_k \preceq t I\), we can always choose \(t\) to be big enough such that the strict inequality holds. For the first matrix constraint, \(w_k = \sum_{j \in N(i)} y_{ij}, \mu = 1 + \varepsilon\) is a strictly feasible solution given that \(\lambda^*_2 \varepsilon > 0\) (Göring et al. 2008). Hence the Slater’s condition holds and strong duality holds between (P2) and (D2).

By Lagrangian duality, one can get the dual problem of (P2):

\[
\begin{align*}
\text{maximize} & \quad tr(\Phi_1) \\
\text{subject to} & \quad tr(\Phi_2) = 1, \\
& \quad tr(11^T \Phi_1) = 0, \\
& \quad tr(E_k \Phi_2) - tr(E_k \Phi_1) = 0, \ k = 1, 2, \ldots, \vert \mathcal{E} \vert, \\
& \quad \Phi_1, \Phi_2 \succeq 0.
\end{align*}
\]

The Karush-Kuhn-Tucker (KKT) conditions for (P2) reads:

\[
\begin{align*}
tr(\Phi_1) & = 1, \\
tr(11^T \Phi_1) & = 0, \\
tr(E_k \Phi_2) - tr(E_k \Phi_1) & = 0, \ k = 1, 2, \ldots, \vert \mathcal{E} \vert, \\
tr(\Phi_2) & = \sum_{k=1}^{\vert \mathcal{E} \vert} y_k^* tr(E_k \Phi_2^*), \quad t^* = \sum_{k=1}^{\vert \mathcal{E} \vert} y_k^* tr(E_k \Phi_2^*),
\end{align*}
\]

with the primal feasibility constraints omitted here for the sake of brevity.

Thunberg & Hu (2016) considered a "topology free" problem, namely, their problem formulation is similar to (LQ1), but without the constraint \(U = (L_w \otimes K)X\). Their control is distributed and turns out to have the form of \(U = (L_w \otimes K)X\) as well, whose edge-weighted Laplacian matrix in the controller corresponds to a complete graph with equal weights \(\frac{1}{N}\) on each edge. Thus their controller is optimal to any other control with a different control gain, graph topologies which is not complete and edge weights not equal to \(\frac{1}{N}\). However, Thunberg & Hu (2016) have not shown the uniqueness of the optimal solution (which can actually be shown using the uniqueness of the solution for the ARE). Here we provide another proof from the view of an optimization problem. Plus, we show the uniqueness of the optimal solution.

**Theorem 6 (c.f. Thunberg & Hu (2016))** Among all controls that belong to the family (5), if given the freedom of choosing positive semidefinite matrix \(Q\) and the graph topology among \(N\) nodes as long as its weighted Laplacian matrix \(L_w \succeq 0\), the optimal energy cost is reached by the control generated by \(Q = 0\) and \(L_w\) corresponding to a complete graph with equal edge weights.

**Proof.** In this case, the matrices \(E_k\) in all should represent the topology of a fully connected graph so that (P2) can have the biggest feasible domain. Choose \(y_k^* = \frac{1}{N}, \ k = 1, \ldots, \vert \mathcal{E} \vert, \ t^* = 1\) and

\[
\Phi_1^* = \Phi_2^* = \begin{bmatrix}
\frac{1}{N} & -\frac{1}{N(N-1)} & \cdots & -\frac{1}{N(N-1)} \\
-\frac{1}{N(N-1)} & \frac{1}{N} & \cdots & \cdots \\
\vdots & \vdots & \ddots & \vdots \\
-\frac{1}{N(N-1)} & \cdots & -\frac{1}{N(N-1)} & \frac{1}{N}
\end{bmatrix},
\]

which is feasible to the dual problem (D2). All KKT conditions are then trivially fulfilled, except for the complementarity slackness. Plugging in \(t^* = 1\) and \(y_k^* = \frac{1}{N}\), the complementarity slack conditions read as

\[
\begin{align*}
tr([I - y_0 11^T - \sum_{k=1}^{\vert \mathcal{E} \vert} y_k^* E_k]\Phi_1^*) & = 1 - \frac{1}{N} \sum_{k=1}^{\vert \mathcal{E} \vert} tr(E_k \Phi_1^*) = 1 - \frac{1}{N} tr(L_w^T \Phi_1^*) = 0, \\
tr([\sum_{k=1}^{\vert \mathcal{E} \vert} y_k^* E_k - t^* I]\Phi_2^*) & = \frac{1}{N} \sum_{k=1}^{\vert \mathcal{E} \vert} tr(E_k \Phi_2^*) - 1 \\
& = \frac{1}{N} tr(L_w^T \Phi_2^*) - 1 = 0.
\end{align*}
\]
where $L^t$ denotes the Laplacian matrix of a fully-connected unweighted graph. Since $\Phi_1^* = \Phi_2^*$, (16) is the same as (17). And we know that $tr(\Phi_1^*) = 1$ and $tr(11^T\Phi_1^*) = 0$, then \( \frac{1}{N} tr[(L^t + 11^T)\Phi_1^*] = 1 \). Therefore, the KKT conditions are satisfied and $y_k^* = \frac{1}{N}$ is an optimal solution to the problem (P2).

Now we show the uniqueness of the solution. Suppose apart from $y_k^* = \frac{1}{N}$, $\forall k$, there is another optimal solution $y_k^*$ also reaches the optimal value $t^* = 1$. Both $y_k^*$ and $y_k^*$ would have to satisfy the constraints, namely:

\[
I - y_0^* 11^T \preceq \sum_{k=1}^{n(n-1)} y_k^* E_k \preceq I \tag{18}
\]
\[
I - y_0^* 11^T \preceq \sum_{k=1}^{n(n-1)} y_k^* E_k \preceq I \tag{19}
\]

By subtracting (18) and (19), we have

\[
\sum_{k=1}^{n(n-1)} (y_k^* - y_k^*) E_k \preceq y_0^* 11^T.
\]

This means for any $v \perp 1$, it must hold

\[
\sum_{k=1}^{n(n-1)} (y_k^* - y_k^*) v^T E_k v \leq 0,
\]
\[
\sum_{k=1}^{n(n-1)} (y_k^* - y_k^*) v^T E_k v \leq 0.
\]

Since $E_k \succeq 0$, $\forall k$, then it must hold $y_k^* = y_k^*$.

Note that since the optimal value of the problem (P1) is 1, this means that $\lambda_2 = \lambda_3 = \cdots = \lambda_N$. Recall that the energy cost function is $\lambda_1$, this means that $y_k^*$ is the same as $y_k^*$ and we have the following theorem:

**Theorem 8** Suppose $\Phi_1^*$ and $\Phi_2^*$ are the optimal solutions to the dual problem (D2) for the graph $G(V, E, W)$. Denote $\hat{G}(V, \hat{E}, \hat{W})$, where $\hat{E} = E \cup \{e_{|\mathbb{E}|+1}\}$, $e_{|\mathbb{E}|+1} \notin \mathbb{E}$, $\hat{W} = W \cup \{w_{|\mathbb{E}|+1}\}$, $w_{|\mathbb{E}|+1} \notin W$. If $tr(E_{|\mathbb{E}|+1}\Phi_2^*) - tr(E_{|\mathbb{E}|+1}\Phi_1^*) = 0$, then there must exist at least one optimal edge-weight realization of $\hat{G}(V, \hat{E}, \hat{W})$ where $w_{|\mathbb{E}|+1} < 0$ and $t^* = \hat{t}^*$.

**PROOF.** Note that $\hat{G}(V, \hat{E}, \hat{W})$ is adding one constraint in the dual. If $tr(E_{|\mathbb{E}|+1}\Phi_2^*) - tr(E_{|\mathbb{E}|+1}\Phi_1^*) = 0$, then the optimal solution $\Phi_1^*$ and $\Phi_2^*$ of the dual problem corresponding to $G(V, E, W)$ should also be an optimal solution to the dual problem corresponding to $\hat{G}(V, \hat{E}, \hat{W})$.

Together with the KKT conditions for the optimal edge-
The KKT conditions for Problem (P3) are

\[
\begin{align*}
tr(\Phi_\ast^1) & = 1, \quad tr(11^T\Phi_\ast^1) = 0, \\
tr(E_k\Phi_\ast^1) - tr(E_k\Phi_\ast^2) & = 0, \quad k = 1, \ldots, |\mathcal{E}|, \\
tr(E_{|\mathcal{E}|+1}\Phi_\ast^1) - tr(E_{|\mathcal{E}|+1}\Phi_\ast^2) & = 0, \\
tr(\Phi_\ast^1) & = \sum_{k=1}^{\mathcal{E}|} y_k^* tr(E_k\Phi_\ast^1), \\
t^* & = \sum_{k=1}^{\mathcal{E}|} y_k^* tr(E_k\Phi_\ast^2),
\end{align*}
\]

(21)

If there exists an optimal edge-weight realization of \(G(V, E, W)\), \(w_{|\mathcal{E}|+1}^* \leq 0\), then the following optimization problem should have the same optimal value \(\hat{t}^*\) as that of (P2), i.e., \(t^* = \hat{t}^*\):

\[
\begin{align*}
\text{minimize} & \quad \hat{t} \\
\text{subject to} & \quad I - \hat{y}\hat{0}11^T \preceq \sum_{k=1}^{\mathcal{E}|} \hat{y}_k E_k \preceq \hat{t}I, \\
& \quad \hat{y}_{|\mathcal{E}|+1} \leq 0.
\end{align*}
\]  
(P3)

The KKT conditions for Problem (P3) are

\[
\begin{align*}
tr(\hat{\Phi}_\ast^1) & = 1, \quad tr(11^T\hat{\Phi}_\ast^1) = 0, \\
tr(E_k\hat{\Phi}_\ast^1) - tr(E_k\hat{\Phi}_\ast^2) & = 0, \quad k = 1, \ldots, |\mathcal{E}|, \\
tr(E_{|\mathcal{E}|+1}\hat{\Phi}_\ast^1) - tr(E_{|\mathcal{E}|+1}\hat{\Phi}_\ast^2) & = 0, \\
\hat{y}_{|\mathcal{E}|+1}^* (tr(E_{|\mathcal{E}|+1}\hat{\Phi}_\ast^1) - tr(E_{|\mathcal{E}|+1}\hat{\Phi}_\ast^2)) & = 0, \\
tr(\hat{\Phi}_\ast^1) & = \sum_{k=1}^{\mathcal{E}|} \hat{y}_k^* tr(E_k\hat{\Phi}_\ast^1) + \hat{y}_{|\mathcal{E}|+1}^* tr(E_{|\mathcal{E}|+1}\hat{\Phi}_\ast^1), \\
t^* & = \hat{t}^* = \sum_{k=1}^{\mathcal{E}|} \hat{y}_k^* tr(E_k\hat{\Phi}_\ast^2) + \hat{y}_{|\mathcal{E}|+1}^* tr(E_{|\mathcal{E}|+1}\hat{\Phi}_\ast^2).
\end{align*}
\]

(22)

Example 9 Consider the graph in Fig. 2.

[Figure 2 is not provided, but it is assumed to be there.] We call the graph without the dashed line \(G(V, E, W)\) and the one with dashed line \(G(V, \hat{E}, \hat{W})\). One optimal edge-weight realization for \(G(V, W, \mathcal{E})\) is \(y_{12} = 2.0620, y_{23} = 0.7215, y_{27} = 1.4877, y_{29} = 0.7667, y_{34} = 2.0620, y_{38} = 1.7297, y_{45} = 1.3747, y_{56} = 1.3747, y_{67} = 2.0620\) with the optimal value \(t^* = 7.2480\).

One optimal edge-weight realization for \(G(V, \hat{E}, \hat{W})\) is \(\hat{y}_{12} = 2.0620, \hat{y}_{23} = 1.0825, \hat{y}_{27} = 1.0827, \hat{y}_{28} = 0.8457, \hat{y}_{34} = 2.0620, \hat{y}_{38} = 0.9147, \hat{y}_{45} = 1.3747, \hat{y}_{56} = 1.3747, \hat{y}_{67} = 2.0620, \hat{y}_{78} = 0.9143\) with all edge-weight positive and same optimal value \(\hat{t}^* = t^* = 7.2480\) as that of \(G(V, \hat{E}, \hat{W})\).

Another optimal edge-weight realization of \(G(V, \hat{E}, \hat{W})\) is \(y_{12} = 2.0620, y_{23} = 0.6602, y_{27} = 1.5079, y_{28} = 0.8404, y_{34} = 2.0620, y_{38} = 1.8680, y_{45} = 1.3747, y_{56} = 1.3747, y_{67} = 2.0620, y_{78} = -0.4566\). In this optimal edge-weight realization, agent 7 and 8 are “competing with” or “push away from” each other, but this will not effect the system performance. □

Another sufficient condition for the existence of negative edge weights is the following:

Theorem 10 Suppose \(\Phi_\ast^1\) and \(\Phi_\ast^2\) are the optimal solutions to the dual problem (D2) for the graph \(G(V, E, W)\). If \(tr(E_{|\mathcal{E}|+1}\Phi_\ast^2) - tr(E_{|\mathcal{E}|+1}\Phi_\ast^1) > 0\), then the optimal solution to the optimal weight realization of \(G(V, \hat{E}, \hat{W})\) must satisfy \(w_{|\mathcal{E}|+1}^* \leq 0\) and \(t^* > \hat{t}^*\).

Proof. To prove this, we first construct the following optimization problem:

\[
\begin{align*}
\text{minimize} & \quad \hat{t} \\
\text{subject to} & \quad I - \hat{y}\hat{0}11^T \preceq \sum_{k=1}^{\mathcal{E}|} \hat{y}_k E_k \preceq \hat{t}I, \\
& \quad \hat{y}_{|\mathcal{E}|+1} \geq 0.
\end{align*}
\]  
(P4)
Example 11 Consider the graph in Fig. 3.

We call the graph without the dashed line $\hat{G}(V, \hat{W}, \hat{E})$ and the one with dashed line $G(V, W, E)$. One optimal edge-weight realization for $G(V, W, E)$ is $y_{13} = 0.5329$, $y_{16} = 0.4896$, $y_{18} = 0.5675$, $y_{23} = 0.7777$, $y_{27} = 0.5072$, $y_{34} = 0.1534$, $y_{35} = 0.4235$, $y_{36} = 0.4331$, $y_{45} = 0.7028$, $y_{47} = 0.5991$, $y_{56} = 0.7604$, $y_{68} = 0.4663$, $y_{68} = 0.8208$ with the optimal value $t^* = 3.0592$.

tr$(E_{|E|+1}\hat{P}_2^2) - tr(E_{|E|+1}\hat{P}_1^2) = 0$, (24)

while by assumption it holds that $tr(E_{|E|+1}\hat{P}_2^4) - tr(E_{|E|+1}\hat{P}_1^4) > 0$. Since $(\{y_k\}, y_{|E|+1} = 0)$ is the optimal solution to (P4), this means $(\{y_k\}, y_{|E|+1} = 0)$ is the best edge weight realization among all $(\{y_k\}, y_{|E|+1} \geq 0)$. But it is not the optimal edge weight realization to $G(V, \hat{E}, W)$, which means for $\hat{G}(V, \hat{E}, W)$, it must satisfy $w^*_{|E|+1} < 0$ and $t^* > \hat{t}^*$. □

Remark 12 From the two theorems above, it can be seen that if we already have an optimal solution of the original graph and we want to add an edge between two nodes, we can use the sign of $tr(E_{|E|+1}\hat{P}_2^4) - tr(E_{|E|+1}\hat{P}_1^4)$ to determine if the two agents will "compete with" each other or not before doing the computation. Moreover, if $tr(E_{|E|+1}\hat{P}_2^4) - tr(E_{|E|+1}\hat{P}_1^4) < 0$, then the optimal solution to the optimal edge weight realization of $\hat{G}(V, \hat{E}, W)$ must satisfy $w^*_{|E|+1} > 0$ and $t^* > \hat{t}^*$, which can be shown in a similar way as Theorem 10.

4.4 Distributed Optimization in the Case of Regular Graphs

When the graph is $\kappa$-regular, namely, each node has $\kappa$ neighbours, this also means $|E(i)| = \kappa$; each node is able to solve the problem (P2) in a distributed manner. In fact, this case is quite common in the real applications since usually for homogeneous nodes that construct the network, each node has the same amount of communication channel. This assumption will be used later and explained in Remark 15.

4.4.1 Intersection of Convex Sets

In order to develop the distributed optimization algorithm for regular graphs, we first rewrite the dual prob-
Therefore the idea of this upcoming distributed optimization problem (D2) into a standard SDP form:

\[
\text{maximize } \quad tr(I\Phi) \\
\text{subject to } \quad tr(I\Phi) = 1, \quad tr(1\Phi) = 0, \quad tr(\bar{E}_k\Phi) = 0, \quad k = 1, 2, \cdots, |\mathcal{E}|, \\
\Phi \succeq 0.
\]

where \( \Phi = \text{diag}(\Phi_1, \Phi_2), \bar{I} = \text{diag}(I_N, 0), \bar{I} = \text{diag}(0, I_N), I = \text{diag}(11^T, 0) \) and \( \bar{E}_k = \text{diag}(-E_k, E_k) \).

From the node’s perspective, we can rewrite the feasible domain \( \Omega \) of (DS) as an intersection of node’s local feasible domain:

\[
\Omega = \{ \Phi \mid tr(I\Phi) = 1, \quad tr(1\Phi) = 0, \quad \Phi \succeq 0, \quad tr(\bar{E}_k\Phi) = 0, \quad k = 1, 2, \cdots, |\mathcal{E}|, \}
\]

\[
= \bigcap_{i \in \mathcal{V}} \Omega_i,
\]

where

\[
\Omega_i = \{ \Phi^{(i)} \mid tr(I\Phi^{(i)}) = 1, \quad tr(1\Phi^{(i)}) = 0, \quad \Phi^{(i)} \succeq 0, \quad tr(\bar{E}_k\Phi^{(i)}) = 0, \quad k \in \mathcal{E}(i) \}.
\]

Therefore the idea of this up-coming distributed optimization set-up is that each node solves its own \( \Phi^{(i)*} \) and all \( \Phi^{(i)*} \) reach consensus in the intersection of the \( \Omega_i \). Hence a penalty term on consensus of \( \Phi^{(i)} \) is added to the objective function and leads to the following problem

\[
\text{minimize } \quad \sum_{i \in \mathcal{V}} \left\{ -\frac{1}{N} tr(I\Phi^{(i)}) + \frac{M}{2} \sum_{j \in N(i)} \|\Phi^{(i)} - \Phi^{(j)}\|_F^2 \right\}
\]

subject to \( \Phi^{(i)} \in \Omega_i, \)

where \( M \) is the penalty parameter.

\textbf{Proposition 13} Problem (26) has the following solution \( \{ \Phi^{(1)*}, \cdots, \Phi^{(N)*} \} \), where \( \Phi^{(i)*} \approx \Phi^{(j)*}, \forall i \neq j \) and \( \Phi^{(i)*} \) is almost the optimal solution to (DS).

\textbf{Proof.} The KKT conditions of (26) read

\[
tr(I\Phi^{(i)*}) = 1, \quad tr(1\Phi^{(i)*}) = 0, \quad \Phi^{(i)*} \succeq 0, \\
tr(\bar{E}_k\Phi^{(i)*}) = 0, \quad k \in \mathcal{E}(i), \quad S^{(i)*} \succeq 0, \\
- \frac{1}{N} \bar{I} + t^{(i)*} \bar{I} + y^{(i)*} 1 - \sum_{k \in \mathcal{E}(i)} y^{(i)*}_k \bar{E}_k - S^{(i)*} \\
+ 2M \sum_{j \in N(i)} (\Phi^{(j)*} - \Phi^{(j)*}) = 0, \quad (27)
\]

\[
tr(S^{(i)*}\Phi^{(i)*}) = 0, \quad \forall i \in \mathcal{V},
\]

Because of the penalty term in (26), \( \Phi^{(i)*} \approx \Phi^{(j)*}, \forall i \neq j \), then summing-up the second last equations in (27) for all \( i \in \mathcal{V} \) will result in "almost" the same KKT conditions of (DS).

\textbf{Remark 14} We would like to use Primal Dual Interior Point Method (PDIPM) (see Alizadeh et al., 1998) to solve the SDP problem because it is a robust and efficient method for SDP problems. The reason we use a penalty term on consensus of \( \Phi^{(i)} \) instead of using consensus constraint \( \sum_{j \in N(i)} (\Phi^{(j)} - \Phi^{(j)}) = 0 \) is that PDIPM uses Newton iteration to solve the perturbed KKT conditions. However, the consensus constraint is introduced, then this also results in the introduction of Lagrangian multiplier terms \( \sum_{j \in N(i)} (Z^{(i)} - Z^{(i)}) \) into the perturbed KKT conditions. Together with the consensus constraint \( \sum_{j \in N(i)} (\Phi^{(j)} - \Phi^{(j)}) = 0 \), a rank deficient matrix \( L_u \otimes I_N \) will show up in the matrix \( Q' \) while equation \( Q' \Delta p^l = r^l \) solves the Newton direction \( \Delta p^l \) of the \( l \)th iteration and hence makes \( Q' \) singular. \( \square \)

\subsection{4.4.2 Primal Dual Interior Point Method}

We solve problem (26) using PDIPM. If we are not satisfied with result on the consensus among \( \Phi^{(i)*} \), then \( M \) is increased and we use the optimal solution as the initial guess for the new problem, which is called as "warm start". This technique is called Sequential Unconstrained Minimization Technique (SUMT).

Now we solve the problem (26) using PDIPM. The penalized barrier problem of parameter \( \rho > 0 \) reads

\[
\text{minimize } \quad \sum_{i \in \mathcal{V}} \left\{ -\frac{1}{N} tr(I\Phi^{(i)}) - \rho \ln \det \Phi^{(i)} + \frac{M}{2} \sum_{j \in N(i)} \|\Phi^{(i)} - \Phi^{(j)}\|_F^2 \right\}
\]

subject to \( \Phi^{(i)} \in \Omega_i, \)

In order to make the matrix \( Q' \) symmetric in the \( Q' \Delta p^l = r^l \) while computing the Newton step \( \Delta p^l \), the perturbed complementary slackness \( \Phi^{(i)} S^{(i)} = \rho I_{2N} \) is
The next step is to compute the step length \( \alpha \) been proved that the algorithm converge exponentially. where \( \hat{\alpha} \) in \((29)\) and hence we get the equation \( \hat{\alpha} \)

\[
2M \sum_{j \in N(i)} \Delta \phi^{(j)} + \Delta \phi^{(j)} + \Delta y^{(j)} \hat{\alpha} = - \sum_{k \in E(i)} \Delta y^{(k)} \hat{\alpha} + \Delta s^{(k)} = r^{(i)}_{\text{primal}},
\]

where \( r^{(i)}_{\text{primal}} = \frac{1}{N} - 2M \sum_{j \in N(i)} (\phi^{(j)} - \phi^{(i)}) - t^{(i)} \hat{\alpha} - y^{(j)} \hat{\alpha} + \Delta s^{(j)} \). We do symmetric vectorization on both hand sides of \((28)\). Note that when we do the symmetric vectorization on the last equation, it holds that

\[
\Delta \phi^{(j)} \Delta s^{(j)} + \Delta \phi^{(j)} \Delta s^{(j)} + \Delta s^{(j)} \Delta \phi^{(j)} = \rho I_{2n} - \Delta \phi^{(j)} s^{(j)} + s^{(j)} \phi^{(j)}.
\]

Therefore \((28)\) of the \( l \)th iteration can be written in the matrix form:

\[
Q^{(i)} \Delta p^{(i)} = r^{(i)},
\]

where \( Q^{(i)} = \text{diag}(Q^{(i)}), \ldots, Q^{(N)}) + L_{u} \otimes C = [Q^{(1)}T, \ldots, Q^{(N)}T], \Delta p^{(i)} = [\Delta \phi^{(1)}T, \ldots, \Delta \phi^{(N)}T], r^{(i)} = [r^{(1)}, \ldots, r^{(N)}T]. \) (The structure of the detailed matrices can be found in the appendix.)

Remark 15 Although for each agent \( i \), it only knows its own matrix block \( Q^{(i)} \), it does not mean that agent \( i \) is ignorant of the column dimension of the row block \( Q^{(i)} \).

Since the graph is \( \kappa \) regular, each block \( Q^{(i)} \) must have the same dimension. Thus each row block \( Q^{(i)} \) is totally accessible by agent \( i \). \( \Box \)

Gauss elimination can be further done on \( svec(\Delta s^{(i)}) \) in \((29)\) and hence we get the equation \( \hat{\alpha} = r^{(i)} \), where \( \hat{\alpha} = \text{diag}(\hat{\alpha}^{(1)}, \ldots, \hat{\alpha}^{(N)}) + L_{u} \otimes \hat{C} = [\hat{Q}^{(1)}T, \ldots, \hat{Q}^{(N)}T], \Delta p^{(i)} = [\Delta \phi^{(1)}T, \ldots, \Delta \phi^{(N)}T], r^{(i)} = [r^{(1)}, \ldots, r^{(N)}T]. \) Now that every node knows the block \( \hat{\alpha}_{j} \) and \( r^{(i)} \), the approach proposed by Mou et al. \((2015)\) is used to solve the system of linear equations in a distributed manner. By using the algorithm, each node is able to compute its own \( \Delta \phi^{(i)} \) and all \( \Delta \phi^{(j)} \) computed by each node will reach consensus. It has been proved that the algorithm converge exponentially. The next step is to compute the step length \( \alpha \). The step length \( \alpha \) for each node is computed through

\[
\alpha = \max \{ \alpha | \phi^{(i)} + \alpha \Delta \phi^{(i)} \geq 0, \}
\]

\[
S_{0}^{(i)} + \alpha \Delta S_{0}^{(i)} \geq 0 \}\}
\]

Now every node has its own \( \alpha \), but in order to make a correct Newton step, we need an \( \alpha \) such that \( 
\alpha' \Delta \phi^{(i)} + \alpha' \Delta \phi^{(i)} \geq 0, S_{0}^{(i)} + \alpha' \Delta S_{0}^{(i)} \geq 0 \) for all nodes. In order to do this, each node can compare its own \( \alpha \) with those of its neighbours; and choose

\[
\alpha = \min \{ \alpha, \alpha \}, \quad \forall j \in N(i).
\]

This process can be seen as the propagation of the smallest step-length \( \alpha \) from its node to the entire network. The maximum number of steps for every node to get \( \alpha \) is the length of the longest path in the graph.

As a summary, for each node the algorithm can be expressed as:

\textbf{Algorithm 1 Distributed PDIPM}

\textbf{Input:} \( \Phi^{(i)}, \forall i \in \mathcal{V}, \epsilon > 0, \theta (0, 1), tol, \xi > 1 \)

\textbf{While} max \( \| \Phi^{(i)} - \Phi^{(j)} ||_{F} > tol, j \in N(i) \)

\textbf{While} \( \rho > \epsilon \)

\textbf{each node compute} \( \Delta \phi^{(i)} \)

\textbf{compute} \( \alpha \)

\textbf{\( \Phi^{(i)} \leftarrow \Phi^{(i)} + \alpha \Delta \phi^{(i)} \)}

\textbf{\( S^{(i)} \leftarrow S^{(i)} + \alpha \Delta S^{(i)} \)}

\textbf{\( \rho \leftarrow \rho - 1 \)}

\textbf{\( l \leftarrow l + 1 \)}

\textbf{end}

\textbf{\( \Phi^{(i)} \leftarrow \Phi^{(i)} \)}

\textbf{\( M \leftarrow M \xi \)}

\textbf{end}

\textbf{Example 16} Consider the graph in \((4)\). Choose \( M = \)

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{example3.png}
\caption{Example 3}
\end{figure}

\( 500, \epsilon = 10^{-13}, \theta = 0.1, \rho = 10000, \xi = 2 \) and we
get $y_{12}^{(1)*} = y_{12}^{(2)*} = y_{23}^{(3)*} = y_{34}^{(3)*} = y_{34}^{(4)*} = y_{41}^{(4)*} = 0.2505$, $\max \| \Phi(i) - \Phi(j) \| F = 1.0016 \times 10^{-4}$, $j \in N(i), \forall i \in V$. Hence $y_{12}^* = y_{23}^* = y_{34}^* = y_{41}^* = 0.5010$ while the optimal solution of centralized problem (P2) is $y_{12}^* = y_{23}^* = y_{34}^* = y_{41}^* = 0.5000$.

5 Conclusion and Future Work

In this paper, we study consensus control for linear systems with optimal energy cost. Due to the hardness of solving the problem, we focus on a classical family of controllers that are based on ARE and guarantee consensus. A suboptimal energy controller that depends only on the relative information between the agents has been constructed by optimizing the control gain as well as the edge weights of the graph. We have shown that the controller coincides with the results in Thunberg & Hu (2016) when the graph topology is complete and provide two sufficient conditions for the existence of negative optimal edge weights when a new edge is added to the graph. Moreover, we proposed a distributed optimization algorithm for solving the problem when the graph is $\kappa$-regular.

The authors are currently working on improving the convergence rate of the distributed optimization algorithm as well as developing a distributed optimization algorithm for maximizing the algebraic connectivity of graphs by adjusting the edge-weights.

References

Alizadeh, F., Haeberly, J.-P. A., & Overton, M. L. (1998). Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. SIAM Journal on Optimization, 8, 746–768.

Annergren, M., Pakazad, S. K., Hansson, A., & Wahlberg, B. (2014). A distributed primal-dual interior-point method for loosely coupled problems using admm. arXiv preprint arXiv:1406.2192.

Borrelli, F., & Keviczky, T. (2008). Distributed lqr design for identical dynamically decoupled systems. Automatic Control, IEEE Transactions on, 53, 1901–1912.

Cao, Y., & Ren, W. (2010). Optimal linear-consensus algorithms: an lqr perspective. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 40, 819–830.

Dall’Anese, E., Zhu, H., & Giannakis, G. B. (2013). Distributed optimal power flow for smart microgrids. IEEE Transactions on Smart Grid, 4, 1464–1475.

Deshpande, P., Menon, P. P., Edwards, C., & Postlethwaite, I. (2011). A distributed control law with guaranteed lqr cost for identical dynamically coupled linear systems. In Proceedings of the 2011 American Control Conference (pp. 5342–5347). IEEE.

Fax, J. A., & Murray, R. M. (2004). Information flow and cooperative control of vehicle formations. IEEE transactions on automatic control, 49, 1465–1476.

Göring, F., Helmberg, C., & Wappler, M. (2008). Embedded in the shadow of the separator. SIAM Journal on Optimization, 19, 472–501.

Jalili, M. (2013). Enhancing synchronizability of diffusively coupled dynamical networks: a survey. Neural Networks and Learning Systems, IEEE Transactions on, 24, 1009–1022.

Lin, F., Fardad, M., & Jovanovic, M. R. (2011). Augmented lagrangian approach to design of structured optimal state feedback gains. IEEE Transactions on Automatic Control, 56, 2923–2929.

Lin, F., Fardad, M., & Jovanovic, M. R. (2013). Design of optimal sparse feedback gains via the alternating direction method of multipliers. IEEE Transactions on Automatic Control, 58, 2426–2431.

Lou, Y., Shi, G., Johansson, K. H., & Hong, Y. (2014). Approximate projected consensus for convex intersection computation: Convergence analysis and critical error angle. IEEE Transactions on Automatic Control, 59, 1722–1736.

Mou, S., Liu, J., & Morse, A. S. (2015). A distributed algorithm for solving a linear algebraic equation. IEEE Transactions on Automatic Control, 60, 2863–2878.

Nedic, A., & Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. IEEE Transactions on Automatic Control, 54, 48–61.

Nedic, A., Ozdaglar, A., & Parrilo, P. A. (2010). Constrained consensus and optimization in multi-agent networks. IEEE Transactions on Automatic Control, 55, 922–938.

Pakazad, S. K., Hansson, A., Andersen, M. S., & Rantzer, A. (2014). Distributed robustness analysis of interconnected uncertain systems using chordal decomposition. IFAC Proceedings Volumes, 47, 2594–2599.

Pakazad, S. K., Hansson, A., Andersen, M. S., & Rantzer, A. (2015). Distributed semidefinite programming with application to large-scale system analysis. arXiv preprint arXiv:1504.07755.

Rogge, J., Suykens, J., & Aeyels, D. (2010). Consensus over ring networks as a quadratic optimal control problem. IFAC Proceedings Volumes, 43, 317–323.

Schäcke, K. (2013). On the kronecker product. Simonetto, A., & Leus, G. (2014). Distributed maximum likelihood sensor network localization. IEEE Transactions on Signal Processing, 62, 1424–1437.

Thunberg, J., & Hu, X. (2016). Optimal output consensus for linear systems: a topology free approach. Automatica, 68, 352–356.

Trentelman, H., Stoorvogel, A. A., & Hautus, M. (2012). Control theory for linear systems. Springer Science & Business Media.

Willems, J. (1971). Least squares stationary optimal control and the algebraic riccati equation. IEEE Transactions on Automatic Control, 16, 621–634.

Yi, P., Hong, Y., & Liu, F. (2015). Distributed gradi-
ent algorithm for constrained optimization with application to load sharing in power systems. Systems & Control Letters, 83, 45–52.
Zhang, H., Lewis, F. L., & Das, A. (2011). Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback. IEEE Transactions on Automatic Control, 56, 1948–1952.

A The Structure of the Matrices in Newton Iteration

$$\Delta p^{(i)T} = [\Delta y_{0}^{(i)}, \Delta \Phi^{(i)^{T}}, \Delta y_{k_{1}}^{(i)}, \cdots, \Delta y_{k_{N}}^{(i)}, \text{svec}(\Delta \Phi^{(i)^{T}})^{T},$$

$$\text{svec}(\Delta S^{(i)l})]^{T} = [\Delta p_{\text{primal}}^{(i)^{T}}, \Delta p_{\text{dual}}^{(i)^{T}}, \Delta p_{\text{comp}}^{(i)^{T}},$$

where $k_{1}, \cdots, k_{N} \in \mathcal{E}(i)$.

$$Q^{(i)} = \begin{bmatrix}
0_{N+2} & A_{i} & 0_{(N+2)\times(2N+1)} \\
A_{i}^{T} & 0_{(N+2\times(2N+1)} & -I_{N(2N+1)} \\
0_{(N+2\times(2N+1)} & S^{(i)l} \otimes_s I_{2N} & \Phi^{(i)^{l}} \otimes_s I_{2N}
\end{bmatrix},$$

$$C = \begin{bmatrix}
0_{N+2} & 0_{(N+2)\times(2N+1)} & 0_{(N+2\times(2N+1)} \\
0_{(N+2\times(2N+1)} & 2MI_{N(2N+1)} & 0_{(N+2\times(2N+1)} \\
0_{(N+2\times(2N+1)} & 0_{(N+2\times(2N+1)} & 0_{(N+2\times(2N+1)}
\end{bmatrix},$$

$$A_{i}^{T} = [\text{svec}(I), \text{svec}(I), \text{svec}(E_{k_{1}}), \cdots, \text{svec}(E_{k_{N}})],$$

where $k_{1}, \cdots, k_{N} \in \mathcal{E}(i)$.

$$r^{(i)^{T}} = [-tr(1\Phi^{(i)^{l}}), 1 - tr(1\Phi^{(i)^{l}}), tr(1\Phi^{(i)^{l}}), \cdots, \text{tr}(1\Phi^{(i)^{l}}), r^{(i)^{T}}_{\text{primal}}, S^{(i)l} - \frac{1}{2}(\Phi^{(i)^{l}}S^{(i)l} + S^{(i)l}\Phi^{(i)^{l}})^{T}]$$

$$= [r_{\text{dual}}^{(i)^{T}}, r_{\text{primal}}^{(i)^{T}}, r_{\text{comp}}^{(i)^{T}},$$

where $k_{1}, \cdots, k_{N} \in \mathcal{E}(i)$.

$$\hat{Q}^{(i)} = \begin{bmatrix}
0_{N+2} & A_{i} \\
A_{i}^{T} (\Phi^{(i)^{l}} \otimes_s I_{N})^{-1}(S^{(i)l} \otimes_s I_{N})
\end{bmatrix},$$

$$\hat{C} = \begin{bmatrix}
0_{N+2} & 0_{(N+2)\times(2N+1)} \\
0_{(N+2\times(2N+1)} & 2MI_{N(2N+1)}
\end{bmatrix},$$

$$\hat{p}^{(i)^{T}} = [\Delta p_{\text{primal}}^{(i)^{T}}, \Delta p_{\text{dual}}^{(i)^{T}},$$

$$\hat{r}^{(i)^{T}} = [r_{\text{dual}}^{(i)^{T}}, r_{\text{primal}}^{(i)^{T}} + r_{\text{comp}}^{(i)^{T}}(\Phi^{(i)^{l}} \otimes_s I_{N})^{-T}].$$