Utility of Edge-wise Funnel Coupling for Asymptotically Solving Distributed Consensus Optimization

Jin Gyu Lee*, Thomas Berger**, Stephan Trenn***, and Hyungbo Shim*

***Abstract*— A new approach to distributed consensus optimization is studied in this paper. The cost function to be minimized is a sum of local cost functions which are not necessarily convex as long as their sum is convex. This benefit is obtained from a recent observation that, with a large gain in the diffusive coupling, heterogeneous multi-agent systems behave like a single dynamics whose vector field is simply the average of all agents' vector fields. However, design of the large coupling gain requires global information such as network structure and individual agent dynamics. In this paper, we employ a nonlinear time-varying coupling of diffusive type, which we call 'edge-wise funnel coupling.' This idea is borrowed from adaptive control study, which enables decentralized design of distributed optimizer without knowledge on global information. Interestingly, without a common internal model, each agent achieves asymptotic consensus to the optimal solution of the global cost. We illustrate this result by a network that asymptotically finds the least-squares solution of a linear equation in a distributed manner.**

I. Introduction

Recent developments in the fields such as formation control, smart grid, and resilient state estimation have posed a question of how to design a network so that agents collectively find an optimizer [1]–[5], and consensus optimization is a vast research field which is a subclass of the aforementioned problem. Let the cost function be given as

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = \sum_{i=1}^{N} f_i(x), \tag{1}$$

which is the sum of N heterogeneous cost functions. The question of how to construct a dynamic system for each node $i \in \mathcal{N} := \{1, \ldots, N\}$ that finds the minimizer $x^* \in \mathbb{R}^n$ of $f(\cdot)$, with each node i having access to its individual cost function $f_i(\cdot)$ only, has been tackled in recent years [6]–[12]. However, most of them, e.g., [6]–[11], assume that the individual cost function $f_i(\cdot)$ is convex. This is the outcome of the need of stability, e.g., passivity, for each node, to achieve consensus.

In this paper, we present a network that finds the minimizer x^* of $f(\cdot)$ asymptotically, with the assumption that $f(\cdot)$ is strictly convex even if each function $f_i(\cdot)$ is not necessarily convex. This is obtained from a recent observation that, with a large coupling gain in the diffusive coupling, heterogeneous multi-agent systems behave like a single dynamics whose vector field is simply the average of all agents' vector fields [13], [14]. By this observation, it is possible to trade stability among agents, and therefore, to relax assumptions on the individual cost function. However, there are some limitations, for instance

1) it only guarantees practical consensus, and
2) design of the coupling gain, that is used for each agent, requires global information such as network structure and individual agent dynamics.

To resolve these issues, we modify the linear diffusive term of the designed network into a nonlinear time-varying coupling, which we call 'edge-wise funnel coupling.' This idea is motivated by the funnel control studied in adaptive controls [15].

It is emphasized that we obtain *asymptotic* consensus to the unique minimizer x^* of $f(\cdot)$ by the proposed funnel coupling. This, in fact, seems to violate the common presumption, in the funnel control community, that asymptotic tracking of an arbitrary reference signal with prescribed performance is not possible. A recent discovery in [16] presented a trick how to overcome this restriction, and finally achieved asymptotic tracking via a funnel gain. This paper borrows the same technique from [16] to achieve asymptotic consensus (without additional dynamics like the PI consensus algorithms nor embedding a common internal model).

The organization of the rest of the paper is as follows. In Section II, we detail our problem formulation and also introduce a network that is designed with a constant coupling gain. In Section III, the given network is modified with the motivation given from funnel control, in the purpose of resolving limitations in the former design. Section IV illustrates the utility of this design by an example of a distributed least-squares solver. Finally, Section V concludes the paper.

Notation: Laplacian matrix $\mathcal{L} = [l_{ij}] \in \mathbb{R}^{N\times N}$ of a graph is defined as $\mathcal{L} := \mathcal{D} - \mathcal{A}$, where $\mathcal{A} = [\alpha_{ij}]$ is the adjacency matrix of the graph and $\mathcal{D}$ is the diagonal matrix with its i-th diagonal entry being $\sum_{j=1}^{N} \alpha_{ij}$. By its construction, it contains at least one eigenvalue of zero, whose corresponding eigenvector is $1_N := [1, \ldots, 1]^T \in \mathbb{R}^N$, and all the other eigenvalues have non-negative real parts. For undirected graphs, the zero eigenvalue is simple if and only if the corresponding graph is connected. For vectors or matrices a

This work was partially supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) via the grant BE 6263/1-1, by the National Research Foundation of Korea grant funded by the Korea government(MSIP) (No. 2015R1A2A2A01003878) and by the NWO vidi grant 639.032.733.

*ljgman@cdsl.kr, hshim@snu.ac.kr, ASRI, Department of Electrical and Computer Engineering, Seoul National University, Korea

**thomas.berger@uni-hamburg.de, Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany

***s.trenn@rug.nl, Bernoulli Institute, University of Groningen, Netherlands

and b, $\mathrm{col}(a,b) := [a^T, b^T]^T$. The operation defined by the symbol $\otimes$ is the Kronecker product. The maximum norm of a vector x is defined by $\|x\|_\infty := \max_i |x_i|$, and the Euclidean norm is denoted by $\|x\| := \sqrt{x^T x}$. The induced maximum norm of a matrix A is written by $\|A\|_\infty$. The gradient of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is defined as $\partial f := \mathrm{col}(\partial f/\partial x^1, \dots, \partial f/\partial x^n)$. The identity matrix of size $m \times m$ is denoted by I_m.

II. Problem Setting and Preliminaries

Consider a network of N agents where each agent $i \in \mathcal{N} = \{1, \dots, N\}$ has access to its own cost function $f_i : \mathbb{R}^n \to \mathbb{R}$ but not others f_j, $j \neq i$. Here, $f_i(\cdot)$ satisfies the following.

Assumption 1: For each $i \in \mathcal{N}$, $f_i(\cdot)$ is continuously differentiable, and its gradient $\partial f_i(\cdot)$ is globally Lipschitz with a Lipschitz constant L_i, i.e., $\|\partial f_i(x) - \partial f_i(x')\| \le L_i \|x - x'\|$ for all $x, x' \in \mathbb{R}^n$. //

The network objective is to solve, in a distributed way,

$$\mathrm{minimize}_x \quad f(x) = \sum_{i=1}^{N} f_i(x)$$

under the following assumption.

Assumption 2: The sum of N cost functions

$$f(x) = \sum_{i=1}^{N} f_i(x)$$

is strictly convex, i.e.,

$$f(tx + (1-t)x') < tf(x) + (1-t)f(x'),$$

for any $t \in (0,1)$ and $x, x' \in \mathbb{R}^n$ such that $x \neq x'$. Moreover, there exists a point $x^* \in \mathbb{R}^n$ such that $f(x^*) \le f(x)$ for all $x \in \mathbb{R}^n$. //

Since the function $f(\cdot)$ is strictly convex, there exists unique minimizer $x^* \in \mathbb{R}^n$, and $\partial f(x)$ becomes zero only at x^*. Therefore, the gradient descent algorithm given by

$$\dot{\hat{x}} = -\partial f(\hat{x}) = -\sum_{i=1}^{N} \partial f_i(\hat{x}) \quad \in \mathbb{R}^n \tag{2}$$

solves the optimization problem. In particular, the solution $\hat{x}$ asymptotically converges to the unique minimizer.

Motivated by this, we can think of a distributed algorithm, in which individual dynamics of agent $i \in \mathcal{N}$ is given by

$$\dot{x}_i = -\partial f_i(x_i) + k \sum_{j \in \mathcal{N}_i} (x_j - x_i) \quad \in \mathbb{R}^n \tag{3}$$

where $k > 0$ is a design parameter, and $\mathcal{N}_i$ is a subset of $\mathcal{N}$ whose elements are the indices of the connected agents that send information to the agent i.

Remark 1: Insight behind the proposed network (3) comes from the so-called 'blended dynamics' approach [13], [14]. In this approach, the behavior of heterogeneous multi-agent systems

$$\dot{x}_i = g_i(t, x_i) + k \sum_{j \in \mathcal{N}_i} (x_j - x_i), \quad i \in \mathcal{N},$$

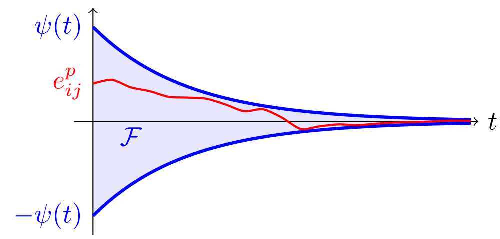

Fig. 1. The funnel: a pre-designed time-varying error bound

with large coupling gain k is approximately estimated for the time interval $[0, \infty)$ by the behavior of the blended dynamics defined by

$$\dot{\hat{x}} = \frac{1}{N} \sum_{i=1}^{N} g_i(t, \hat{x})$$

under the assumption that the blended dynamics is stable. In our case, the blended dynamics is obtained as

$$\dot{\hat{x}} = -\frac{1}{N} \sum_{i=1}^{N} \partial f_i(\hat{x}) = -\frac{1}{N} \partial f(\hat{x})$$

which then becomes the (scaled) gradient descent algorithm given in (2). //

Assumption 3: The graph is undirected and connected. //

Proposition 1: Let Assumptions 1, 2, and 3 hold. Then, for any compact set $K \subset \mathbb{R}^{Nn}$, and for any $\eta > 0$, there exists $k^* > 0$ such that, for each $k > k^*$ and $\mathrm{col}(x_1(0), \dots, x_N(0)) \in K$, the solution to (3) exists for all $t \ge 0$, and satisfies

$$\limsup_{t \to \infty} \|x_i(t) - x^*\| \le \eta, \quad \forall i \in \mathcal{N}. \quad //$$

Proof: See [14]. ■

Although this result is already quite powerful, it still has a disadvantage that the optimizer is not found asymptotically but just approximately. Moreover, for computing the threshold k^*, global information such as the network topology and all f_i's is needed, and so, the method is not completely decentralized. These drawbacks will be resolved in the next section by choosing the gain k adaptively based on the idea of funnel control.

III. Edge-wise Funnel Coupling

Building on the idea of the edge-wise funnel coupling law [17], we propose to replace the static diffusive coupling term $k \sum_{j \in \mathcal{N}_i} (x_j - x_i)$ of agent i by the following coupling law

$$\sum_{j \in \mathcal{N}_i} K\left(\frac{x_j - x_i}{\psi(t)}\right) \cdot \frac{x_j - x_i}{\psi(t)}$$

where $\psi : \mathbb{R}_{\ge 0} \to \mathbb{R}_{>0}$ is the so-called funnel boundary function (Figure 1), and $K : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ is defined by

$$K(\eta) := \mathrm{diag}\left(\frac{1}{1 - |\eta^1|}, \cdots, \frac{1}{1 - |\eta^n|}\right).$$

By introducing $e_{ij} := x_j - x_i$, the dynamics of agent i now become

$$\dot{x}_i = -\partial f_i(x_i) + \sum_{j\in\mathcal{N}_i} \mathrm{col}\left(\frac{e_{ij}^1}{\psi(t)-|e_{ij}^1|}, \dots, \frac{e_{ij}^n}{\psi(t)-|e_{ij}^n|}\right) \quad (4)$$

where $x_i = \mathrm{col}(x_i^1, \dots, x_i^n)$ and $e_{ij}^p = x_j^p - x_i^p$.

Intuition of the funnel coupling in (4) is simple. If the p-th component of the difference between two agents, $e_{ij}^p(t) = x_j^p(t) - x_i^p(t)$, approaches the funnel boundary $\pm\psi(t)$ so that $\psi(t) - |e_{ij}^p(t)|$ becomes closer to zero, then the gain to $e_{ij}^p(t)$ gets larger towards infinity. Therefore, if there is only one neighbor, then the state x_i tends to its neighbor x_j since the large coupling term dominates the vector field $-\partial f_i(x_i)$, and the error $e_{ij}^p(t)$ remains inside the funnel. However, with more than one neighbor, this intuition becomes no longer straightforward because two neighbors may attract x_i in opposite direction with almost infinite power. Actual analysis shows that all the errors $e_{ij}(t)$ remain inside the funnel, which is however far more complicated. In this paper, we simply quote one of the main results in [18] as follows.

Proposition 2: Let Assumptions 1, 2, and 3 hold. Then, for any bounded continuously differentiable function $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ with bounded derivative, and for any initial conditions $x_i(0)$ with $\|x_j(0) - x_i(0)\|_\infty < \psi(0)$ for all $j \in \mathcal{N}_i$, $i \in \mathcal{N}$, the solution to (4) exists for all $t \geq 0$ and satisfies

$$\|x_j(t) - x_i(t)\|_\infty < \psi(t), \quad \forall t \geq 0, \quad \forall i \in \mathcal{N},\ j \in \mathcal{N}_i.$$

Moreover, if there exists $\overline{M}$ such that $\|x_i(t)\| \leq \overline{M}$ for all $t \geq 0$, then there exists $\epsilon > 0$ such that

$$\frac{\|x_j(t) - x_i(t)\|_\infty}{\psi(t)} \leq 1 - \epsilon, \quad \forall i \in \mathcal{N},\ j \in \mathcal{N}_i,$$

for all $t \geq 0$. //

Now according to Proposition 2, if we select $\psi(\cdot)$ such that $\lim_{t\to\infty} \psi(t) = 0$, then we obtain asymptotic consensus, i.e., $\lim_{t\to\infty} \|x_j(t) - x_i(t)\|_\infty = 0$. This in turn implies that, with a definition of a new variable $x_{\text{avg}} := (1/N)\sum_{i=1}^N x_i$, each state $x_i(t)$ tends to $x_{\text{avg}}(t)$ as time gets large. Now, by Assumption 3, it is seen that

$$\sum_{i=1}^N \sum_{j\in\mathcal{N}_i} \frac{e_{ij}^p}{\psi(t) - |e_{ij}^p(t)|} = 0$$

for $p = 1, \dots, n$. Therefore, we have that

$$\dot{x}_{\text{avg}} = -\frac{1}{N}\sum_{i=1}^N \partial f_i(x_i) \quad \to \quad -\frac{1}{N}\sum_{i=1}^N \partial f_i(x_{\text{avg}})$$

as $t \to \infty$, and so, it is rather intuitive that the coupled system (4) will asymptotically find the unique minimizer x^*. This intutition is made precise by the following theorem, which is our main result.

Theorem 1: Let Assumptions 1, 2, and 3 hold. Then, for any bounded continuously differentiable function $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ with bounded derivative and satisfying $\lim_{t\to\infty} \psi(t) = 0$, and for any initial conditions $x_i(0)$ with $\|x_j(0) - x_i(0)\|_\infty < \psi(0)$ for all $j \in \mathcal{N}_i$, $i \in \mathcal{N}$, the solution to (4) exists for all $t \geq 0$ and satisfies

$$\lim_{t\to\infty} x_i(t) = x^*, \quad \forall i \in \mathcal{N},$$

i.e. each agent's state converges to the global optimizer. Furthermore, there exists $\epsilon > 0$ such that

$$\frac{\|x_j(t) - x_i(t)\|_\infty}{\psi(t)} \leq 1 - \epsilon, \quad \forall i \in \mathcal{N},\ j \in \mathcal{N}_i,$$

for all $t \geq 0$, i.e. the coupling gain K remains bounded. //

Proof: Let L_i denote a Lipschitz constant of ∂f_i, in Assumption 1, and let $\mathcal{T}$ be an arbitrary spanning tree in the graph with the corresponding incidence matrix $T \in \mathbb{R}^{N\times(N-1)}$. From the fact that

$$\frac{1}{N} 1_N 1_N^T + T(T^T T)^{-1} T^T = I_N,$$

let $\mathfrak{t}_i^T$ be the i-th row of $T(T^T T)^{-1}$. Then, with $x_{\text{avg}} = (1/N)\sum_{i=1}^N x_i$ and $\tilde{x} := (T^T \otimes I_n)\mathrm{col}(x_1, \dots, x_N)$, we have $x_i = x_{\text{avg}} + (\mathfrak{t}_i^T \otimes I_n)\tilde{x}$. Therefore, it follows that

$$\dot{x}_{\text{avg}} = -\frac{1}{N}\sum_{i=1}^N \partial f_i(x_{\text{avg}} + (\mathfrak{t}_i^T \otimes I_n)\tilde{x}).$$

Note that by Proposition 2, we have

$$\|(T^T \otimes I_n)\mathrm{col}(x_1, \dots, x_N)\|_\infty = \|\tilde{x}\|_\infty < \psi(t), \quad \forall t \geq 0.$$

Now, let $V(x_{\text{avg}}) := f(x_{\text{avg}}) - f(x^*) = \sum_{i=1}^N (f_i(x_{\text{avg}}) - f_i(x^*))$. Then, the function V is positive definite, and thus, we have class $\mathcal{K}$ functions $\alpha_1(\cdot)$ and $\alpha_2(\cdot)$ such that

$$\alpha_1(\|x - x^*\|) \leq V(x) \leq \alpha_2(\|x - x^*\|), \quad \forall x \in \mathbb{R}^n.$$

Now, the derivative of V satisfies

$$\begin{aligned}
\dot{V} &\leq \partial f(x_{\text{avg}})^T \dot{x}_{\text{avg}} \\
&= -\frac{1}{N}\partial f(x_{\text{avg}})^T \sum_{i=1}^N \partial f_i(x_{\text{avg}} + (\mathfrak{t}_i^T \otimes I_n)\tilde{x}) \\
&= -\frac{1}{N}\|\partial f(x_{\text{avg}})\|^2 \\
&\quad - \frac{1}{N}\partial f(x_{\text{avg}})^T \sum_{i=1}^N [\partial f_i(x_{\text{avg}} + (\mathfrak{t}_i^T \otimes I_n)\tilde{x}) - \partial f_i(x_{\text{avg}})] \\
&\leq -\frac{1}{N}\|\partial f(x_{\text{avg}})\|^2 + \frac{1}{N}\|\partial f(x_{\text{avg}})\| \sum_{i=1}^N L_i \|(\mathfrak{t}_i^T \otimes I_n)\tilde{x}\| \\
&\leq -\frac{1}{N}\|\partial f(x_{\text{avg}})\| \left(\|\partial f(x_{\text{avg}})\| - L^*\psi(t)\right),
\end{aligned}$$

where $L^* := \sqrt{n}\sum_{i=1}^N L_i \|T(T^T T)^{-1}\|_\infty$.

Since $f(\cdot)$ is strictly convex, $\|\partial f(x)\|$ equals zero only when $x = x^*$. Therefore, there exits class $\mathcal{K}$ function $\alpha_3(\cdot)$ such that, for any $x \in \mathbb{R}^n$ we have

$$\|\partial f(x)\| \geq \alpha_3(\|x - x^*\|).$$

Now, for any $\eta > 0$, there exists $T_\eta > 0$ such that

$$\psi(t) < \frac{\eta}{2L^*}, \quad \forall t \geq T_\eta.$$

Then, for all $t \geq T_\eta$, we obtain $\dot{V} \leq -(1/N)(\eta^2/2) < 0$ whenever $V \geq \alpha_2(\alpha_3^{-1}(\eta))$. This implies that

$$\limsup_{t\to\infty} \|x_{\text{avg}}(t) - x^*\| \leq \limsup_{t\to\infty} \alpha_1^{-1}(V(t)) \leq \alpha_1^{-1}(\alpha_2(\alpha_3^{-1}(\eta))).$$

However, the selection of η was arbitrary, so that we have $\lim_{t\to\infty} \|x_{\text{avg}}(t) - x^*\| = 0$. Then, the results of Proposition 2 conclude the proof. ■

Remark 2: The asymptotic convergence result of Theorem 1 may seem contradictory to a common presumption in the funnel control community that asymptotic tracking of an arbitrary reference signal with prescribed performance is not possible. However, a recent paper [16] has overcome this restriction by a relatively simple trick to rewrite the funnel control law. We inherit the same technique in this paper. Indeed, the coupling gain

$$\frac{1}{\psi(t) - |x_j^p - x_i^p|}$$

grows unbounded when asymptotic consensus is achieved because $\psi(t) \to 0$ as $t \to \infty$ and this implies $\psi(t) - |x_j^p - x_i^p|$ also tends to zero. However, simply rewriting the coupling term as

$$\frac{1}{1 - |x_j^p - x_i^p|/\psi(t)} \frac{x_j^p - x_i^p}{\psi(t)}$$

we see that by the result of Theorem 1, the fraction $|x_j^p - x_i^p|/\psi(t)$ is bounded away from 1, hence the new gain and the total input are bounded even if $1/\psi(t)$ tends to infinity.

Emphasis is also made that the result of Theorem 1 may seem to violate another presumption in the synchronization research that heterogeneous multi-agent systems can not asymptotically synchronize without a common internal model. This violation is resolved by observing that we use a time-varying coupling law, which is not considered in the framework of the internal model principle for multi-agent systems [19].

Finally, we address that the difference between asymptotic consensus versus practical consensus may not be very important in practical applications, as long as the residual error in practical consensus is sufficiently small. In this sense, our concern on asymptotic convergence may be of more academic interest. //

IV. Example: Distributed Least-Squares Solver

As distributed algorithms have been developed in various fields of study so as to divide a large computational problem into small-scale computations, finding the least-squares solution of a given large linear equation in a distributed manner has been tackled in recent years [20]–[23]. Let the equation be given by

$$Ax = b \in \mathbb{R}^M \tag{5}$$

where $A \in \mathbb{R}^{M\times n}$ is a matrix of full column rank, $x \in \mathbb{R}^n$, and $b \in \mathbb{R}^M$. Throughout the section, we suppose that the total of M lines of equations are grouped into N equation banks, and the i-th equation bank consists of m_i lines of equations so that $\sum_{i=1}^N m_i = M$. In particular, we write the i-th equation bank as

$$A_i x = b_i \in \mathbb{R}^{m_i}, \quad i = 1, 2, \ldots, N,$$

where $A_i \in \mathbb{R}^{m_i\times n}$ is the i-th block row of the matrix A, and $b_i \in \mathbb{R}^{m_i}$ is the i-th block element of b.

Finding the least-squares solution x^* of (5) even when $b \notin \text{im}(A)$ can be cast as a simple optimization problem

$$\text{minimize}_x \ \frac{1}{2}\|Ax - b\|^2 = \sum_{i=1}^N \frac{1}{2}\|A_i x - b_i\|^2.$$

By letting $f_i(x) = (1/2)\|A_i x - b_i\|^2$, the problem becomes consensus optimization. Then, according to the recipe in the previous section, we can find the least-squares solution asymptotically by a network given by

$$\begin{aligned} \dot{x}_i &= -A_i^T(A_i x - b_i) \\ &+ \sum_{j\in\mathcal{N}_i} \text{col}\left(\frac{x_j^1 - x_i^1}{\psi(t) - |x_j^1 - x_i^1|}, \ldots, \frac{x_j^n - x_i^n}{\psi(t) - |x_j^n - x_i^n|}\right), \end{aligned} \tag{6}$$

where each agent i uses the information of A_i and b_i only.

Now, for a linear equation given by

$$Ax = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 10 \\ 20 \\ 18 \\ 100 \end{bmatrix} = b$$

with $N = 5$ and each equation bank consisting of a single equation, the gradient descent algorithm

$$\dot{\hat{x}} = -A^T(A\hat{x} - b), \quad \hat{x}(0) = 0,$$

results in convergence of its state to the unique minimizer $x^* = 17$. On the other hand, as guaranteed in Theorem 1, the solution of (6) also converges to the minimizer $x^* = 17$. Indeed, Figure 2.(b) shows a simulation result when the funnel boundary function is chosen as $\psi(t) = \exp(-0.8t)$, the graph is set to a linear graph as in Figure 3, and the initial conditions are set as $x_1(0) = 0$, $x_2(0) = 0.1$, $x_3(0) = -0.1$, $x_4(0) = 0.2$, and $x_5(0) = -0.2$.

For comparison, Figure 2.(c) shows the trajectory of the network

$$\dot{x}_i = -A_i^T(A_i x_i - b_i) + k \sum_{j\in\mathcal{N}_i}(x_j - x_i), \quad i \in \mathcal{N}, \tag{7}$$

with the constant coupling gain $k = 100$. Two figures of Figures 2.(b) and 2.(c) clearly show that the network with the constant coupling gain can only achieve practical convergence to the minimizer $x^* = 17$, while asymptotic convergence is obtained by using edge-wise funnel coupling.

We also inspect the derivative of each $x_i(t)$ because the right-hand side of $\dot{x}_i$ can be considered as an input to each agent, and we are interested whether their values are too large or not. Figure 4 shows $\dot{x}_i(t)$ used for the asymptotic convergence with edge-wide funnel coupling in (6). On the other hand, Figure 5 is for the case of constant coupling

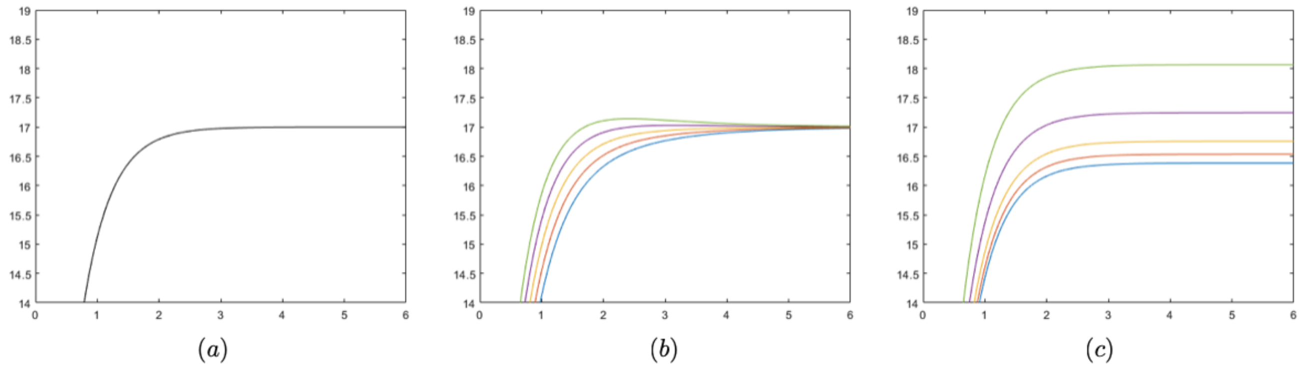

Fig. 2. Solution trajectory of (a) the blended dynamics, (b) the network with edge-wise funnel coupling, and (c) the network with constant coupling gain

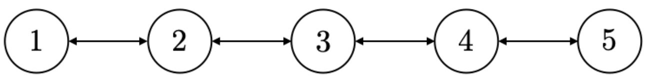

Fig. 3. Underlying graph among five agents

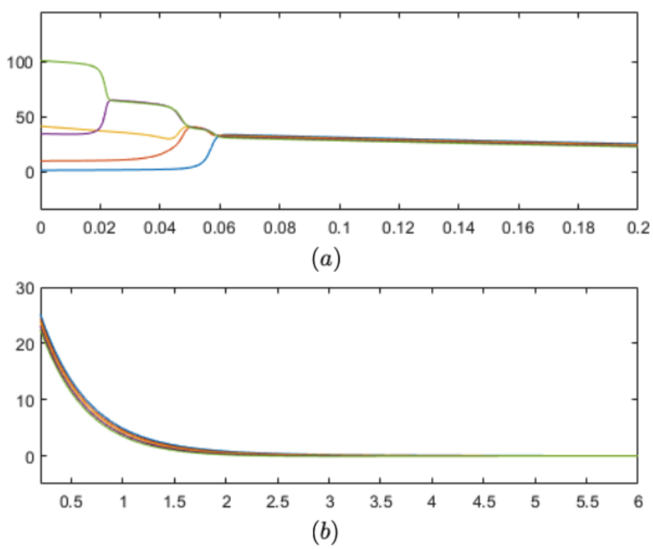

Fig. 4. Plot of $\dot{x}_i(t)$ for the network with edge-wise funnel coupling: (a) $t \in [0, 0.2]$ and (b) $t \in [0.2, 6]$

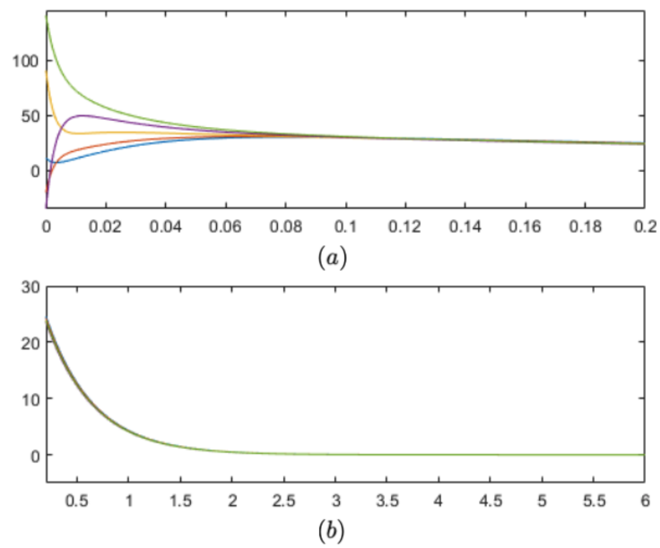

Fig. 5. Plot of $\dot{x}_i(t)$ for the network with constant coupling gain: (a) $t \in [0, 0.2]$ and (b) $t \in [0.2, 6]$

gain in (7). It is verified that their magnitudes are not very different. It is noted that, for the case of funnel coupling, $\dot{x}_i(t)$ is bounded even though the funnel $\psi(t)$ approaches zero. This is because the funnel coupling law

$$\frac{x_j - x_i}{\psi(t) - |x_j - x_i|}$$

can be re-written as

$$\frac{1}{1 - |x_j - x_i|/\psi(t)} \frac{x_j - x_i}{\psi(t)}$$

where $|x_j - x_i|/\psi(t)$ is proven to be bounded away from 1 in Theorem 1.

Finally, it is observed that the diffusive coupling term $(x_j - x_i)/(\psi(t) - |x_j - x_i|)$ converges to a specific constant e_{ij}^* that cancels the heterogeneity of the individual vector field such that

$$-A_i^T(A_i x^* - b_i) + \sum_{j \in \mathcal{N}_i} e_{ij}^* = 0, \quad i \in \mathcal{N}.$$

This is clearly indicated in Figure 4, where we observe that $\dot{x}_i(t)$ converges to zero. By this, we interpret that the edge-wise funnel coupling law still adapts in some sense. Moreover, in this special case, we can even compute e_{ij}^* through the five equations:

$$\begin{aligned} -(17-1) + e_{12}^* &= 0 \\ -(17-10) + e_{23}^* + e_{21}^* &= 0 \\ -2(34-20) + e_{34}^* + e_{32}^* &= 0 \\ -2(34-18) + e_{45}^* + e_{43}^* &= 0 \\ -(17-100) + e_{54}^* &= 0 \end{aligned}$$

with four variables, which yields a unique solution

$$e_{12}^* = 16, \quad e_{23}^* = 23, \quad e_{34}^* = 51, \quad e_{45}^* = 83.$$

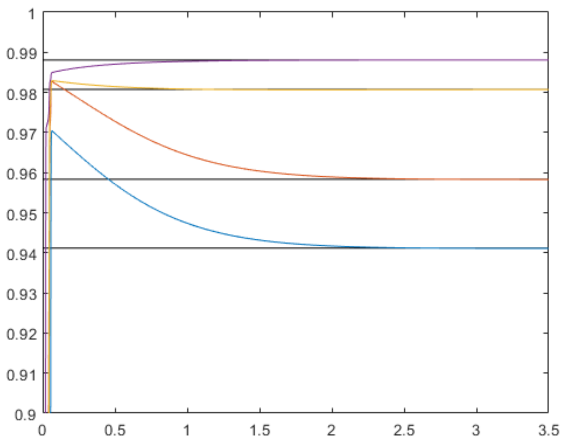

Fig. 6. Trajectory of the fraction $(x_j - x_i)/\psi(t)$

This is also shown in Figure 6, which depicts the convergence of the fraction $(x_j(t) - x_i(t))/\psi(t)$. In particular, if we denote $\eta_{ij}^* := \lim_{t\to\infty}(x_j(t) - x_i(t))/\psi(t)$, then we have

$$\frac{\eta_{ij}^*}{1 - |\eta_{ij}^*|} = e_{ij}^*, \quad \forall i \in \mathcal{N},\ j \in \mathcal{N}_i,$$

which gives finite values of

$$\eta_{12}^* = \frac{16}{17}, \quad \eta_{23}^* = \frac{23}{24}, \quad \eta_{34}^* = \frac{51}{52}, \quad \eta_{45}^* = \frac{83}{84}.$$

V. Conclusion

Based on the design philosophy of the blended dynamics induced by large coupling gain, a network is first designed to solve a distributed consensus optimization by a constant coupling gain. Then, to overcome the limitation of the constant gain design, the network is modified by introducing the edge-wise funnel coupling, whose intuition is borrowed from adaptive control. As a consequence, we obtain a network that achieves asymptotic convergence to the unique minimizer, which doesn't require any global information. The utility of the proposed network is shown by an optimization problem of finding the least-squares solution.

References

[1] A. Nedić and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[2] I. Lobel, A. Ozdaglar, and D. Feijer, "Distributed multi-agent optimization with state-dependent communication," *Mathematical Programming*, vol. 129, no. 2, pp. 255–284, 2011.

[3] J. G. Lee, J. Kim, and H. Shim, "Fully distributed resilient state estimation based on distributed median solver," under review for *IEEE Transactions on Automatic Control*, 2019.

[4] F. Wirth, S. Stuedli, J. Y. Yu, M. Corless, and R. Shorten, "Nonhomogeneous place-dependent Markov chains, unsynchronised AIMD, and network utility maximization," *arXiv preprint arXiv: 1404.5064*, 2014.

[5] M. Corless, C. King, R. Shorten, and F. Wirth, *AIMD dynamics and distributed resource allocation*. SIAM, 2016.

[6] G. Qu and N. Li, "Harnessing smoothness to accelerate distributed optimization," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1245–1260, 2018.

[7] J. Wang and N. Elia, "Control approach to distributed optimization," in *Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing*, 2010, pp. 557–561.

[8] T. Hatanaka, N. Chopra, T. Ishizaki, and N. Li, "Passivity-based distributed optimization with communication delays using PI consensus algorithm," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4421–4428, 2018.

[9] B. Gharesifard and J. Cortés, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781–786, 2014.

[10] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *Proceedings of the 50th IEEE Conference on Decision and Control*, 2011, pp. 3800–3805.

[11] G. Chen and Z. Li, "A fixed-time convergent algorithm for distributed convex optimization in multi-agent systems," *Automatica*, vol. 95, pp. 539–543, 2018.

[12] Z. Li, Z. Ding, J. Sun, and Z. Li, "Distributed adaptive convex optimization on directed graphs via continuous-time algorithms," *IEEE Transactions on Automatic Control*, vol. 63, no. 5, pp. 1434–1441, 2018.

[13] J. Kim, J. Yang, H. Shim, J.-S. Kim, and J. H. Seo, "Robustness of synchronization of heterogeneous agents by strong coupling and a large number of agents," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 3096–3102, 2016.

[14] J. G. Lee and H. Shim, "A tool for analysis and synthesis of heterogeneous multi-agent systems under rank-deficient coupling," under review for *Automatica*, available at `arXiv:1804.00638`, 2019.

[15] A. Ilchmann, E. P. Ryan, and C. J. Sangwin, "Tracking with prescribed transient behaviour," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 7, pp. 471–493, 2002.

[16] J. G. Lee and S. Trenn, "Asymptotic tracking via funnel control," under review for *IEEE Conference on Decision and Control*, 2019.

[17] S. Trenn, "Edge-wise funnel synchronization," in *Proceedings in Applied Mathematics and Mechanics*, vol. 17, pp. 821–822, 2017.

[18] J. G. Lee, T. Berger, S. Trenn, and H. Shim, "Synchronization with prescribed transient behavior: Heterogeneous multi-agent systems under edge-wise funnel coupling," 2019, in preparation.

[19] P. Wieland, J. Wu, and F. Allgöwer, "On synchronous steady states and internal models of diffusively coupled systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2591–2602, 2013.

[20] X. Wang, J. Zhou, S. Mou, and M. J. Corless, "A distributed linear equation solver for least square solutions," in *Proceedings of the 56th IEEE Conference on Decision and Control*, 2017, pp. 5955–5960.

[21] G. Shi and B. D. O. Anderson, "Distributed network flows solving linear algebraic equations," in *Proceedings of the American Control Conference*, 2016, pp. 2864–2869.

[22] G. Shi, B. D. O. Anderson, and U. Helmke, "Network flows that solve linear equations," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2659–2674, 2017.

[23] Y. Liu, Y. Lou, B. D. O. Anderson, and G. Shi, "Network flows as least squares solvers for linear equations," in *Proceedings of 56th IEEE Conference on Decision and Control*, 2017, pp. 1046–1051.