Analysis and design of a distributed k-winners-take-all model ⋆

Yinyan Zhang [a], Shuai Li [b], Bin Xu [c], Yong Yang [d]

[a] *College of Cyber Security, Jinan University, Guangzhou, China*

[b] *College of Engineering, Swansea University, Swansea, United Kingdom*

[c] *School of Automation, Northwestern Polytechnical University, Xi'an, China*

[d] *School of Mechatronic Engineering, Guangdong Polytechnic Normal University, Guangzhou, China*

Abstract

The k-winners-take-all (kWTA) problem is to find the k largest inputs from N inputs. In this paper, we design and propose a novel distributed kWTA model, for which no central unit is needed to realize the computation of the k winners. As a result, the proposed model has the general advantages of distributed models over centralized ones, such as better robustness to faults of agents. The global asymptotic convergence of the proposed distributed model is proven. Besides, two numerical examples on networks of agents with static inputs and time-varying inputs are presented to validate the performance of the proposed model.

Key words: Optimization; k-winners-take-all; Convergence; Multi-agent system; Invariance principle.

1 Introduction

The winners-take-all (WTA) phenomenon refers to the selection of the largest input (called the winner) from given inputs, by which the winner tends to be activated while the others are deactivated (Binas *et al.*, 2014; Li *et al.*, 2013, 2017). As an extension of WTA, in kWTA, the k largest inputs are selected from N inputs (Costea and Marinov, 2011). In the past decades, several kWTA models have been designed and analyzed. Marinov and Calvert (2003) theoretically analyzed the computable restrictions on the parameters of an analog Hopfield-type neural network model that can generate the k-WTA behavior for the case with successive lists of real numbers at a given rate. By formulating the kWTA problem as linear and quadratic programming problems, via the Karush-Kuhn-Tucker (KKT) conditions, Liu and Wang (2008) proposed two kWTA network models activated by discontinuous activation functions. Based on the formulations, continuous-time kWTA models can be derived as recurrent neural network models (Hu and Wang, 2008; Hu and Zhang, 2009; Li *et al.*, 2013; Liu *et al.*, 2010; Liu and Wang, 2006; Wang, 2010; Xia and Sun, 2009), although with different structure complexities. In particular, the kWTA models proposed in Li *et al.* (2013); Liu *et al.* (2010); Liu and Wang (2006); Wang (2010); Xiao *et al.* (2012) are convergent in finite time, and the model in Wang (2010) has only a single variable. For the kWTA model proposed in Wang (2010), it is assumed that the threshold logic units can be implemented. For the situation that such units cannot be perfectly implemented, the theoretical analysis on the performance of the kWTA model was presented in Feng *et al.* (2015). Meanwhile, the robustness analysis for the kWTA model with input noises was performed in Feng *et al.* (2018); Sum, Leung, and Ho (2013). Apart from the continuous-time kWTA model models, a few discrete-time ones were also proposed (Tien, 2017; Tymoshchuk, 2009). It is found that all the aforementioned kWTA models are centralized. In other words, they need a centralized unit to perform all the computation.

During the past two decades, many results on distributed control of multi-agent systems were reported (Jadbabaie *et al.*, 2003; Ni *et al.*, 2017; Olfati-Saber and Murray, 2004; Wen *et al.*, 2017; Yu *et al.*, 2013; Zhang and Li, 2018; Zhang *et al.*, 2019; Zhu *et al.*, 2017). In particular, Olfati-Saber and Murray (2004) discuss several consensus problems for networks of dynamic agents with

⋆ This work is supported by the National Natural Science Foundation of China (61933010) and the Fundamental Research Funds for the Central Universities (3102019ZDHKY13). This paper was not presented at any IFAC meeting. Corresponding author S. Li.

Email addresses: yyzhang@jnu.edu.cn (Yinyan Zhang), shuai.li@swansea.ac.uk (Shuai Li), binxu@gmail.com (Bin Xu), yy2008@gpnu.edu.cn (Yong Yang).

fixed or switching topologies. Recently, kWTA models were applied to the distributed dynamic task allocation of robots (Jin and Li, 2018; Jin *et al.*, 2019), where low-pass average consensus filters are used to estimate the centralized terms in the kWTA models, while the theoretical analysis on the performance of the approaches was performed on the centralized kWTA models. To our knowledge, there is no existing distributed kWTA model with theoretically guaranteed performance.

In this paper, based on the existing results, we propose a novel distributed kWTA model. One of the potential applications of the proposed model is the distributed task allocation. As a preliminary step, we consider the case that the network topology is undirected connected. Our analysis shows that the model has an invariant set, at which the model degrades into an existing centralized kWTA model with global asymptotic convergence. As a result, the performance of the proposed distributed kWTA model is theoretically guaranteed. The difference of the current work from Jin and Li (2018); Jin *et al.* (2019) is two-fold: 1) The asymptotic convergence of the distributed kWTA model proposed in this paper is theoretically proven, while in Jin and Li (2018); Jin *et al.* (2019) theoretical results are not given about the corresponding distributed kWTA model; 2) In the current work, we adopt a high-pass consensus filter which has much faster response to the inputs compared with the low-pass one adopted in Jin and Li (2018); Jin *et al.* (2019). The contributions of this paper mainly include the following: 1) A novel distributed kWTA model is proposed with theoretically proven global asymptotic convergence; 2) To some extent, this work provides some insights about how to develop a distributed model based on an existing centralized dynamic model via the LaSalle's invariance principle and consensus approaches.

2 Preliminary and problem formulation

In this section, some relevant existing results are presented and the problem investigated in this paper is described.

2.1 Centralized kWTA

Mathematically, the kWTA can be described as follows (Hu and Wang, 2008; Hu and Zhang, 2009; Li *et al.*, 2013; Liu *et al.*, 2010; Liu and Wang, 2006; Wang, 2010; Xia and Sun, 2009):

$$z_i = \begin{cases} 1, \text{ if } u_i \in \{k \text{ largest elements of } \mathbf{u}\}, \\ 0, \text{ otherwise}, \end{cases} \tag{1}$$

where $\mathbf{z} = [z_1, z_2, \cdots, z_N]^{\mathrm{T}}$ and $\mathbf{u} = [u_1, u_2, \cdots, u_N]^{\mathrm{T}}$ denote the output and input vector with the dimension being N, respectively. If $z_i = 1$, element i is called a winner; if $z_i = 0$, element i is called a loser.

Let $\mathbf{1}$ denote a $N-$dimensional column vector with each element being 1. According to Theorem 4 of Liu and Wang (2006), we have the following lemma.

Lemma 1: The solution to the following quadratic programming problem with $\mathbf{z}$ being the decision variable is the same as that to (1):

$$\begin{aligned} &\min_{\mathbf{z}\in\mathbb{R}^N} \ a\mathbf{z}^{\mathrm{T}}\mathbf{z} - \mathbf{u}^{\mathrm{T}}\mathbf{z} \\ &\text{subject to } \mathbf{1}^{\mathrm{T}}\mathbf{z} = k, \\ &\qquad z_i \in [0,1], \ \forall i \in \{1, 2, \cdots, N\}, \end{aligned} \tag{2}$$

given that $0 < 2a \leq \tilde{u}_k - \tilde{u}_{k+1}$ with $\tilde{u}_k$ and $\tilde{u}_{k+1}$ denoting the kth largest element and the $k+1$th largest element in $\mathbf{u}$, respectively.

According to Theorem 1 of Xia and Sun (2009), we have the following lemma.

Lemma 2: The output of the following continuous-time model is globally asymptotically convergent to the solution to quadratic program (2):

$$\begin{aligned} &\text{state equation } \dot{\chi} = \gamma(-\mathbf{1}^{\mathrm{T}}\mathbf{z} + k), \\ &\text{output equation } \mathbf{z} = g\left(\mathbf{1}\chi + \frac{\mathbf{u}}{a}\right), \end{aligned} \tag{3}$$

with $\chi \in \mathbb{R}$ denoting the state variable, $\mathbf{z} \in \mathbb{R}^N$ denoting the output vector, and $\gamma > 0 \in \mathbb{R}$ being a design parameter, where the array projection operator $g(\cdot)$ is defined as follows:

$$g(x_i) = \begin{cases} 0 \text{ if } x_i < 0, \\ 1 \text{ if } x_i > 1, \\ x_i \text{ otherwise.} \end{cases}$$

By Lemma 1 and Lemma 2, we directly have the following corollary.

Corollary 1: Given that $0 < 2a \leq \tilde{u}_k - \tilde{u}_{k+1}$, the output of (3) is globally asymptotically convergent to the solution to the kWTA problem (1).

It can be easily found that the kWTA model (3) is centralized as the state variable χ needs the information of all the outputs for all the elements in $\mathbf{u}$. Such a property is also possessed by other existing kWTA models, such as those in (Hu and Wang, 2008; Hu and Zhang, 2009; Li *et al.*, 2013; Liu *et al.*, 2010; Liu and Wang, 2006; Wang, 2010). In this paper, the proposed model that will be discussed in the next section is based on model (3) instead of others due to the fact that model (3) is considered to be the simplest model among the existing ones. As seen from (3), the model has only a single state variable χ, which is independent from the scale of the problem.

2.2 Distributed kWTA problem

Consider an undirected connected network of N agents. Each agent has its own output denoted by z_i and input denoted by u_i. The kWTA is that the agents in the network compete with their neighbors such that k winners will be generated for which $z_i = 1$ and the rest will be the losers with $z_i = 0$. As in Li *et al.* (2017), the Laplacian matrix of the corresponding communication graph $G(\mathbb{V}, \mathbb{E}, W)$ of the network is denoted by L, and the set of neighbors of agent i is denoted by $\mathbb{N}(i)$. The kWTA problem investigated in this paper can be described as finding a model to solve the kWTA problem (1) through local interactions.

3 Distributed kWTA model

In this section, the proposed distributed kWTA model together with the corresponding theoretical analysis is presented.

3.1 Model Description

To solve the distributed kWTA problem, we introduce two state variables for each agent, namely x_i and y_i. Let $\mathbf{x} = [x_1, x_2, \cdots, x_N]^{\mathrm{T}}$ and $\mathbf{y} = [y_1, y_2, \cdots, y_N]^{\mathrm{T}}$. The proposed kWTA model is described as follows:

$$\begin{aligned}
&\text{state equation } \epsilon\dot{\mathbf{x}} = C_0(-\mathbf{y} - \mathbf{z} + \frac{k\mathbf{1}}{N}) - C_0 L\mathbf{x},\\
&\text{state equation } \epsilon\dot{\mathbf{y}} = -L(\mathbf{y} + \mathbf{z}),\\
&\text{output equation } \mathbf{z} = g(\mathbf{x} + \frac{\mathbf{u}}{\alpha}),
\end{aligned} \tag{4}$$

where ϵ, C_0, and α are positive constants. In view of (4), the proposed model is totally different from the model in Li *et al.* (2017). In principle, the design of the current model is based on the centralized model in Xia and Sun (2009) instead of the model in Li *et al.* (2017).

Remark 1: For the proposed kWTA model, the state and output of any agent $i \in \mathbb{V}$ satisfies

$$\begin{aligned}
\epsilon\dot{x}_i &= C_0(-y_i - z_i + \frac{k}{N}) - C_0 \sum_{j\in\mathbb{N}(i)} w_{ij}(x_i - x_j),\\
\epsilon\dot{y}_i &= -\sum_{j\in\mathbb{N}(i)} w_{ij}(y_i - y_j) - \sum_{j\in\mathbb{N}(i)} w_{ij}(z_i - z_j),\\
z_i &= g(x_i + \frac{u_i}{\alpha}).
\end{aligned} \tag{5}$$

As seen from (5), the proposed kWTA model is distributed.

3.2 Theoretical analysis

We have the following theorem regarding the convergence of the proposed kWTA model.

Theorem 1: Given that $\mathbf{y}(0) = 0$, $\lambda_{2\min}(L) \geq 7.5C_0$, where $\lambda_{2\min}(L)$ denotes the second smallest eigenvalue of L, and $0 < 2\alpha \leq s_u$ with s_u denoting the minimum separation among the inputs, the output of kWTA model (4) is globally asymptotically convergent to the solution to the kWTA problem (1).

Proof: Let $\varphi(\mathbf{x}) = [\varphi_1(x_1), \varphi_2(x_2), \cdots, \varphi_N(x_N)]^{\mathrm{T}}$ with

$$\varphi_i(x_i) = \frac{\partial g(x_i)}{\partial x_i} \begin{cases} = 0 \text{ if } x_i < 0 \text{ or } x_i > 1,\\ = 1 \text{ if } 0 < x_i < 1,\\ \in [0, 1], \text{ otherwise.}\end{cases}$$

Clearly, with $\|\cdot\|_1$ denoting the 1−norm, we have $\varphi^{\mathrm{T}}(\mathbf{x})\mathbf{1} = \|\varphi(\mathbf{x})\|_1$. Let $\mathrm{diag}(\varphi(\mathbf{x})) = \partial g(\mathbf{x})/\partial\mathbf{x}, \eta_1 = \mathbf{y} + \mathbf{z} - \mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{z}/N \in \mathbb{R}^N, \eta_2 = \mathbf{1}^{\mathrm{T}}\mathbf{z}/N - k/N \in \mathbb{R}$, and $\eta_3 = L\mathbf{x} \in \mathbb{R}^N$. Then, together with (4), we have

$$\begin{aligned}
\epsilon\dot{\mathbf{x}} &= C_0(-\eta_1 - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{z}}{N} + \frac{k\mathbf{1}}{N}) - C_0 L\mathbf{x}\\
&= -C_0\eta_1 - C_0\mathbf{1}\eta_2 - C_0\eta_3.
\end{aligned} \tag{6}$$

With (6) and (4), we further have

$$\begin{aligned}
\dot{\eta}_2 &= \frac{\mathbf{1}^{\mathrm{T}}\dot{\mathbf{z}}}{N}\\
&= \frac{\mathbf{1}^{\mathrm{T}}\mathrm{diag}(\varphi(\mathbf{x}))\dot{\mathbf{x}}}{N}\\
&= \frac{1}{\epsilon N}\varphi^{\mathrm{T}}(\mathbf{x})(-C_0\eta_1 - C_0\mathbf{1}\eta_2 - C_0\eta_3)\\
&= \frac{1}{\epsilon N}(-C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_1 - C_0\|\varphi(\mathbf{x})\|_1\eta_2 - C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_3).
\end{aligned} \tag{7}$$

Meanwhile, for η_1, with I denoting the N−by-N identity matrix, we have

$$\begin{aligned}
\dot{\eta}_1 &= \dot{\mathbf{y}} + \dot{\mathbf{z}} - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}\dot{\mathbf{z}}}{N}\\
&= -\frac{1}{\epsilon}L(\mathbf{y} + \mathbf{z} - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{z}}{N}\\
&\quad + \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N}) + (I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\dot{\mathbf{x}}\\
&= -\frac{1}{\epsilon}L(\mathbf{y} + \mathbf{z} - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{z}}{N}) - \frac{L\mathbf{1}\mathbf{1}^{\mathrm{T}}}{\epsilon N}\\
&\quad + (I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\dot{\mathbf{x}}.
\end{aligned}$$

Since the graph is undirected connected, we have the following property (Jadbabaie *et al.*, 2003):

$$L\mathbf{1} = 0. \tag{8}$$

As a result, together with the definition of η_1, we have

$$\dot{\eta}_1 = -\frac{1}{\epsilon}L\eta_1 + (I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\dot{\mathbf{x}}. \quad (9)$$

Note that $\mathbf{1}^{\mathrm{T}}\eta_1 = \mathbf{1}^{\mathrm{T}}(\mathbf{y}+\mathbf{z}-\mathbf{1}\mathbf{1}^{\mathrm{T}}\mathbf{z}/N) = \mathbf{1}^{\mathrm{T}}\mathbf{y}+\mathbf{1}^{\mathrm{T}}\mathbf{z} - \mathbf{1}^{\mathrm{T}}\mathbf{z} = \mathbf{1}^{\mathrm{T}}\mathbf{y}$ and $\mathbf{1}^{\mathrm{T}}\dot{\mathbf{y}} = -\mathbf{1}^{\mathrm{T}}L(\mathbf{y}+\mathbf{z})/\epsilon = 0$ due to (8). As a result, we have $\mathbf{1}^{\mathrm{T}}\eta_1 \equiv 0$ given that $\mathbf{y}(0) = 0$. Consequently, we have

$$L\eta_1 = (L + C_1\mathbf{1}\mathbf{1}^{\mathrm{T}})\eta_1 = A\eta_1,$$

where $A = L + C_1\mathbf{1}\mathbf{1}^{\mathrm{T}}$ where $C_1 > 0$ is sufficiently large. (In fact, we only need $C_1 \geq \lambda_{2min}(L)/N$). Clearly, the smallest eigenvalue of A satisfies $\lambda_{\min}(A) = \lambda_{2\min}(L) > 0$ (Jadbabaie *et al.*, 2003), where $\lambda_{\min}(\cdot)$ and $\lambda_{2\min}(\cdot)$ denote the smallest and second smallest eigenvalues, respectively. Then, (9) can be rewritten as follows:

$$\dot{\eta}_1 = -\frac{1}{\epsilon}A\eta_1 + (I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\dot{\mathbf{x}}. \quad (10)$$

Let

$$V_1 = \frac{\epsilon\eta_1^{\mathrm{T}}\eta_1}{2}, V_2 = \frac{\epsilon\eta_2^2}{2}, V_3 = \frac{\epsilon\mathbf{x}^{\mathrm{T}}L\mathbf{x}}{2}.$$

Calculating the time derivative of V_1 along the state trajectory of (10), we have

$$\begin{aligned}\dot{V}_1 &= \epsilon\eta_1^{\mathrm{T}}\dot{\eta}_1 \\ &= -\eta_1^{\mathrm{T}}A\eta_1 + \epsilon\eta_1^{\mathrm{T}}(I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\dot{\mathbf{x}}.\end{aligned}$$

Substituting equation (6) into the above equation yields

$$\begin{aligned}\dot{V}_1 = &-\eta_1^{\mathrm{T}}A\eta_1 - C_0\eta_1^{\mathrm{T}}(I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\eta_1 \\ &- C_0\eta_1^{\mathrm{T}}(I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\varphi(\mathbf{x})\eta_2 \\ &- C_0\eta_1^{\mathrm{T}}(I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N})\mathrm{diag}(\varphi(\mathbf{x}))\eta_3 \\ \leq &-\eta_1^{\mathrm{T}}A\eta_1 + C_0\|\eta_1\|_2^2\lambda_{\max}\left(I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N}\right)\|\mathrm{diag}(\varphi(\mathbf{x}))\|_2 \\ &- C_0\eta_1^{\mathrm{T}}\varphi(\mathbf{x})\eta_2 + \frac{C_0\eta_1^{\mathrm{T}}\mathbf{1}\mathbf{1}^{\mathrm{T}}\varphi(\mathbf{x})\eta_2}{N} \\ &+ C_0\|\eta_1\|_2\lambda_{\max}\left(I - \frac{\mathbf{1}\mathbf{1}^{\mathrm{T}}}{N}\right)\|\mathrm{diag}(\varphi(\mathbf{x}))\|_2\|\eta_3\|_2,\end{aligned}$$

where $\|\cdot\|_2$ denotes the 2-norm of a vector or a matrix (for a matrix it is also call the maximum singular value) and $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of a matrix. According to Section A.1.5 of Boyd and Vandenberghe (2004), $\|\mathrm{diag}(\varphi(\mathbf{x}))\|_2 = \sqrt{\lambda_{\max}(\mathrm{diag}^{\mathrm{T}}(\varphi(\mathbf{x}))\mathrm{diag}(\varphi(\mathbf{x})))} = \lambda_{\max}(\varphi(\mathbf{x})) \leq 1$ since $\varphi(\mathbf{x})$ is diagonal and each diagonal element is not larger than 1. Besides, we have $\lambda_{\max}(I - \mathbf{1}\mathbf{1}^{\mathrm{T}}) = 1$. Thus, we have

$$\begin{aligned}\dot{V}_1 \leq &-\eta_1^{\mathrm{T}}A\eta_1 + C_0\|\eta_1\|_2^2 - C_0\eta_1^{\mathrm{T}}\varphi(\mathbf{x})\eta_2 \\ &+ \frac{C_0\eta_1^{\mathrm{T}}\mathbf{1}\mathbf{1}^{\mathrm{T}}\varphi(\mathbf{x})\eta_2}{N} + C_0\|\eta_1\|_2\|\eta_3\|_2.\end{aligned} \quad (11)$$

Note that, by Cauchy-Schwarz inequality,

$$\begin{aligned}\eta_1^{\mathrm{T}}\varphi(\mathbf{x}) &\leq \|\eta_1\|_2\|\varphi(\mathbf{x})\|_2, \\ \eta_1^{\mathrm{T}}\mathbf{1}\mathbf{1}^{\mathrm{T}}\varphi(\mathbf{x})\eta_2 &= (\eta_1^{\mathrm{T}}\mathbf{1})(\mathbf{1}^{\mathrm{T}}\varphi(\mathbf{x}))\eta_2 \\ &\leq (\|\eta_1\|_2\sqrt{N})(\sqrt{N}\|\varphi(\mathbf{x})\|_2)|\eta_2| \\ &= N\|\eta_1\|_2\|\varphi(\mathbf{x})\|_2|\eta_2|,\end{aligned}$$

where $|\cdot|$ denotes the absolute value. Together with inequality (11), we have

$$\begin{aligned}\dot{V}_1 \leq &-\eta_1^{\mathrm{T}}A\eta_1 + C_0\|\eta_1\|_2^2 + C_0|\eta_2|\|\eta_1\|_2\|\varphi(\mathbf{x})\|_2 \\ &+ \frac{C_0}{N}N\|\eta_1\|_2\|\varphi(\mathbf{x})\|_2|\eta_2| + C_0\|\eta_1\|_2\|\eta_3\|_2 \\ = &-\eta_1^{\mathrm{T}}(A - C_0I)\eta_1 + 2C_0\|\eta_1\|_2\|\varphi(\mathbf{x})\|_2|\eta_2| \\ &+ C_0\|\eta_1\|_2\|\eta_3\|_2.\end{aligned}$$

Note that $\|\varphi(\mathbf{x})\|_1 \geq \|\varphi(\mathbf{x})\|_2^2$ since each element of $\varphi(\mathbf{x})$ lies between 0 and 1. Thus, calculating the time derivative of V_2 along the state trajectory of (7), we have

$$\begin{aligned}\dot{V}_2 &= \epsilon\eta_2\dot{\eta}_2 \\ &= \eta_2\frac{1}{N}(-C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_1 - C_0\|\varphi(\mathbf{x})\|_1\eta_2 - C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_3) \\ &\leq \eta_2\frac{1}{N}(-C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_1 - C_0\|\varphi(\mathbf{x})\|_2^2\eta_2 - C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_3) \\ &= \frac{-C_0\|\varphi(\mathbf{x})\|_2^2\eta_2^2}{N} - \frac{C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_1\eta_2}{N} - \frac{C_0\varphi^{\mathrm{T}}(\mathbf{x})\eta_3\eta_2}{N} \\ &\leq \frac{-C_0\|\varphi(\mathbf{x})\|_2^2\eta_2^2}{N} + \frac{C_0|\eta_2|\|\varphi(\mathbf{x})\|_2\|\eta_1\|_2}{N} \\ &\quad + \frac{C_0\|\varphi(\mathbf{x})\|_2\|\eta_3\|_2|\eta_2|}{N}.\end{aligned}$$

For V_3, with (6), we have $\dot{V}_3 = \epsilon\dot{\mathbf{x}}^{\mathrm{T}}L\mathbf{x} = -C_0\eta_1^{\mathrm{T}}L\mathbf{x} - C_0\eta_2\mathbf{1}^{\mathrm{T}}L\mathbf{x} - C_0\|L\mathbf{x}\|_2^2$. Recalling (8) and the symmetricity of L, we have $\mathbf{1}^{\mathrm{T}}L = 0$. Thus, together with the definition of η_3, we further have

$$\begin{aligned}\dot{V}_3 &= -C_0\eta_1^{\mathrm{T}}L\mathbf{x} - C_0\|L\mathbf{x}\|_2^2 \\ &= -C_0\eta_1^{\mathrm{T}}\eta_3 - C_0\|\eta_3\|_2^2 \leq -C_0\|\eta_3\|_2^2 + C_0\|\eta_1\|_2\|\eta_3\|_2.\end{aligned}$$

Consider the Lyapunov candidate function $V = V_1 + NV_2 + V_3$, which is positive definite in light of the definitions of V_1, V_2, and V_3. The time derivative of V along the state trajectory of the system satisfies

$$
\begin{aligned}
\dot{V} &= \dot{V}_1 + N\dot{V}_2 + \dot{V}_3 \\
&\leq -\eta_1^{\mathrm{T}}(A - C_0 I)\eta_1 + 2C_0\|\eta_1\|_2\|\varphi(\mathbf{x})\|_2|\eta_2| \\
&\quad + C_0\|\eta_1\|_2\|\eta_3\|_2 - C_0\|\varphi(\mathbf{x})\|_2^2\eta_2^2 \\
&\quad + C_0|\eta_2|\|\varphi(\mathbf{x})\|_2\|\eta_1\|_2 + C_0\|\varphi(\mathbf{x})\|_2\|\eta_3\|_2|\eta_2| \\
&\quad - C_0\|\eta_3\|_2^2 + C_0\|\eta_1\|_2\|\eta_3\|_2 \\
&\leq C_0\bigg(- \|\eta_1\|_2^2\left(\frac{\lambda_{\min}(A)}{C_0} - 1\right) - (\|\varphi(\mathbf{x})\|_2\eta_2)^2 \\
&\quad - \|\eta_3\|_2^2 + 3\|\eta_1\|_2(\|\varphi(\mathbf{x})\|_2|\eta_2|) + 2\|\eta_1\|_2\|\eta_3\|_2 \\
&\quad + \|\eta_3\|_2(\|\varphi(\mathbf{x})\|_2|\eta_2|)\bigg) \\
&\leq C_0\bigg(- \|\eta_1\|_2^2\left(\frac{\lambda_{\min}(A)}{C_0} - 1\right) - \frac{(\|\varphi(\mathbf{x})\|_2\eta_2)^2}{2} \\
&\quad - \frac{\|\eta_3\|_2^2}{2} + 3\|\eta_1\|_2(\|\varphi(\mathbf{x})\|_2|\eta_2|) + 2\|\eta_1\|_2\|\eta_3\|_2 \\
&\quad + \|\eta_3\|_2(\|\varphi(\mathbf{x})\|_2|\eta_2|)\bigg) \\
&= -C_0\|\eta_1\|_2^2\left(\frac{\lambda_{\min}(A)}{C_0} - 7.5\right) - C_0\left(\frac{3}{\sqrt{2}}\|\eta_1\|_2 \right. \\
&\quad \left. - \frac{\|\varphi(\mathbf{x})\|_2|\eta_2|}{\sqrt{2}}\right)^2 - C_0\left(\sqrt{2}\|\eta_1\|_2 - \frac{1}{\sqrt{2}}\|\eta_3\|_2\right)^2.
\end{aligned}
$$

Given that $\lambda_{\min}(A) = \lambda_{2\min}(L) \geq 7.5C_0$, we have $\dot{V} \leq 0$. Then, let $\dot{V} = 0$, and we have $\eta_1 = 0$, $\eta_3 = 0$, and $\|\varphi(\mathbf{x})\|\eta_2 = 0$. In the invariant set with $\eta_1 = 0$ and $\eta_3 = 0$, equation (6) becomes

$$\epsilon\dot{\mathbf{x}} = -C_0\mathbf{1}\eta_2 = -C_0\mathbf{1}\left(\frac{\mathbf{1}^{\mathrm{T}}\mathbf{z}}{N} - \frac{k}{N}\right).$$

Left multiplying $\mathbf{1}^{\mathrm{T}}/N$ on both sides of the above equation yields

$$\epsilon\dot{\bar{x}} = \epsilon\frac{\mathbf{1}^{\mathrm{T}}\dot{\mathbf{x}}}{N} = -C_0\left(\frac{\mathbf{1}^{\mathrm{T}}\mathbf{z}}{N} - \frac{k}{N}\right) = -\frac{C_0}{N}(1^{\mathrm{T}}\mathbf{z} - k),$$

where $\mathbf{z} = g(\mathbf{x} + \mathbf{u}/\alpha)$, and $\bar{x} = \sum_{i=1}^{N} x_i/N$ denotes the average of the elements in $\mathbf{x}$. Since $\eta_3 = L\mathbf{x} = 0$ in the invariant set, we have $\mathbf{x} = \bar{x}\mathbf{1}$ and, as a result,

$$\mathbf{z} = g(\mathbf{x} + \mathbf{u}/\alpha) = g(\bar{x}\mathbf{1} + \mathbf{u}/\alpha).$$

By summarizing the above results, in the invariant set, we have

$$
\begin{aligned}
\epsilon\dot{\bar{x}} &= -\frac{C_0}{N}(\mathbf{1}^{\mathrm{T}}\mathbf{z} - k), \\
\mathbf{z} &= g(\bar{x}\mathbf{1} + \frac{\mathbf{u}}{\alpha}),
\end{aligned}
\tag{12}
$$

which can be rewritten as (3) with $a = \alpha$, $\chi = \bar{x}$, and $\lambda = C_0/(N\epsilon)$. According to Corollary 1, given that $0 < 2\alpha \leq \tilde{u}_k - \tilde{u}_{k+1}$ with $\tilde{u}_k$ denoting the kth largest element $\tilde{u}_{k+1}$ denoting the $k+1$th largest element in $\mathbf{u}$, the output of (12) is globally asymptotically convergent to the solution of kWTA problem (1). Then, by, LaSalle's invariance principle (Khalil, 2002; Xu *et al.*, 2019), it is further concluded that with $\mathbf{y}(0) = 0$, $\lambda_{2\min}(L) \geq 7.5C_0$, and $0 < 2\alpha \leq \tilde{u}_k - \tilde{u}_{k+1}$, the output of kWTA model (4) is globally asymptotically convergent to the solution to the kWTA problem (1). The proof is complete. □

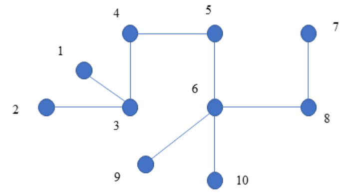

Fig. 1. The communication graph of a network of 10 agents.

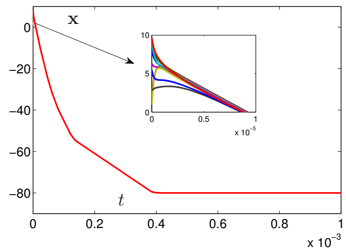

Fig. 2. The transient behavior of state vector $\mathbf{x}$ during the kWTA process with static inputs.

Remark 2: As seen from the proof of Theorem 1, the proof of the convergence of the proposed kWTA is finally converted to that of an existing one thanks to the invariance principle. Essentially, the distributed kWTA is achieved via properly adding additional co-states compared with the centralized kWTA to realize the asymptotic estimation of centralized terms. The design perspective could be useful for obtaining the corresponding distributed models for other problems. By referring to the underlying principle of the design, the proposed model can also be extended to other types of communication topologies that are allowed in the high-pass consensus filter. The detailed theoretical analysis will be further investigated.

Remark 3: In the proposed kWTA model (4), there are three positive parameters ϵ, C_0, and α. As seen from (4), parameter ϵ actually scales the strength of feedback on the state update. In the implementation of the proposed model, to achieve fast convergence, the value of ϵ should be set as small as possible but larger than 0. According to Theorem 1, the value of C_0 need to satisfy $\lambda_{2\min}(L) \geq 7.5C_0$, where $\lambda_{2\min}(L) > 0$ denotes the second smallest eigenvalue of the communication graph. Meanwhile, we also need C_0 to be larger than zero. As a result, the setting of C_0 has similar requirement as ϵ. In practice, the measurement of signals is always with certain resolution due to the limited capability of sensors. For example, the resolution of ultrasonic range finders of

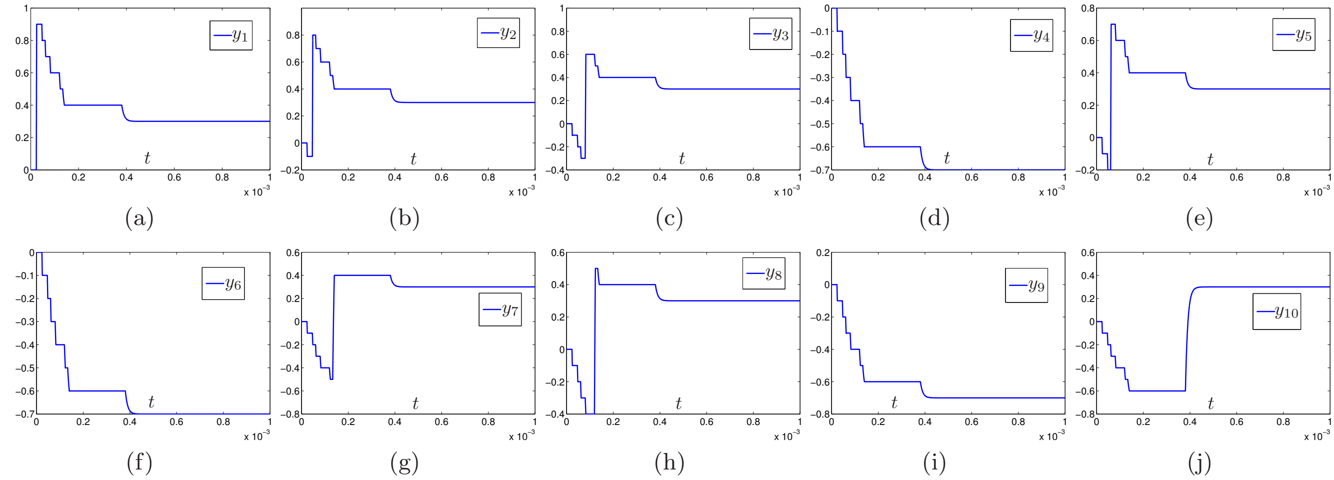

Fig. 3. The transient behavior of state vector $\mathbf{y}$ during the kWTA process with static inputs.

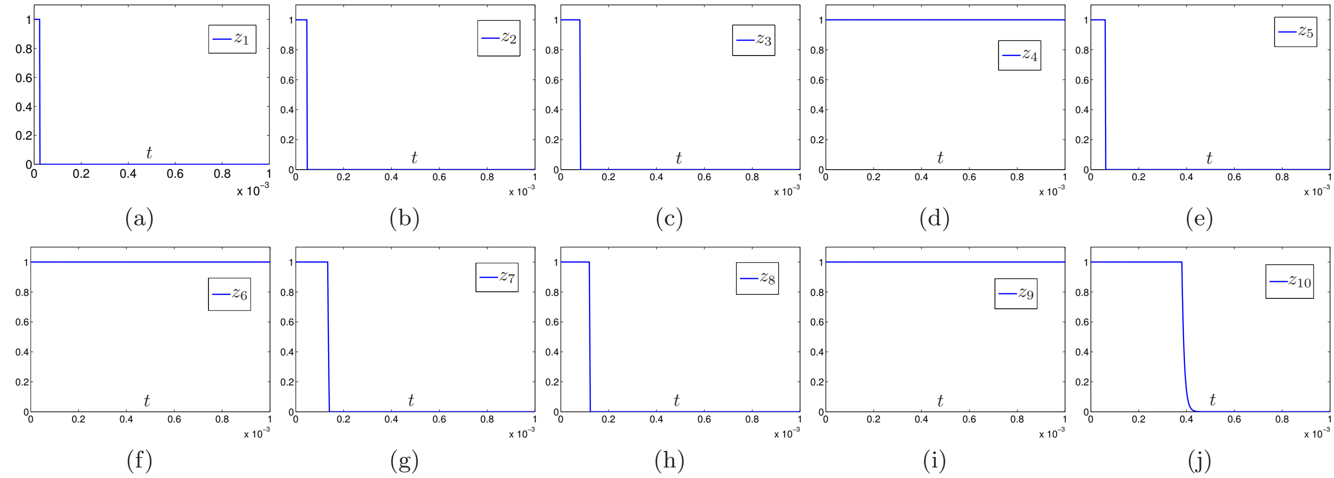

Fig. 4. The transient behavior of output vector $\mathbf{z}$ during the kWTA process with static inputs.

HRLV-MaxSonar-EZ series is 0.001 m. In this case, when the inputs are measurements of such ultrasonic range finders, $s_u = 0.001$. In other words, s_u is determined by the resolution of sensors that generate the inputs. Then, we can set the value of α.

4 Numerical examples

In this section, numerical examples are provided to show the performance of the proposed kWTA model.

4.1 Static inputs

We first consider the kWTA in an undirected connected network of 10 agents, i.e., $N = 10$, with $k = 3$, where the inputs to the agents are static. In other words, the agents need to compete with each other and generate 3 winners. The communication graph is shown in Fig. 1. For convenience of illustration, in the network, each element of the weight matrix W of the graph satisfies $w_{ij} \in \{0, 1\}$. The inputs of the ten agents are static, and they are $\mathbf{u} = [1.1, 2.5, 4, 19, 3.2, 10, 5.5, 5.2, 19.1, 8]^{\mathrm{T}}$. The proposed distributed kWTA model (4) is adopted with the values of the parameters set as $\epsilon = 10^{-9}$, $\alpha = 0.01$, and $C_0 = 0.001$. According to Theorem 1, the initial values of $\mathbf{x}$, i.e., $\mathbf{x}(0)$, are randomly set while $\mathbf{y}(0) = 0$. The simulation results are shown in Fig. 2, Fig. 3, and Fig. 4. It is found that the proposed distributed kWTA model successfully recognizes the k winners.

4.2 Dynamic inputs

We also consider the kWTA in an undirected connected network of 4 agents, i.e., $N = 4$, with $k = 2$, where the inputs to the agents are time-varying. The communication graph is shown in Fig. 5, with the weights labeled beside the edges. The inputs of the 4 agents are given as

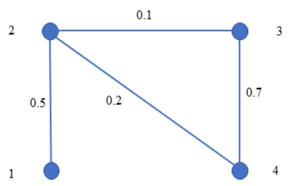

Fig. 5. The communication graph of a network of 4 agents.

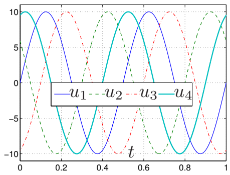

Fig. 6. The transient behavior of time-varying inputs to the agents.

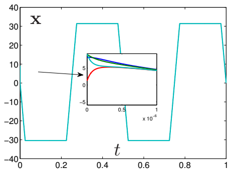

Fig. 7. The transient behavior of state vector $\mathbf{x}$ during the kWTA process with time-varying inputs.

follows:

$$u_i(t) = 10 \sin(4\pi(t + 0.2(i-1)))$$

with $i = 1, 2, 3, 4$, which are plotted on Fig. 6. The proposed distributed kWTA model (4) is adopted with the values of the parameters set as $\epsilon = 10^{-9}$, $\alpha = 0.1$, and $C_0 = 0.01$. The initial state of $\mathbf{x}$ is randomly generated while $\mathbf{y}(0)$ is set to zero according to Theorem 1. As seen from Fig. 7, the state variables x_i reach consensus in almost real time. Meanwhile, Fig. 8 shows that the state variables y_i reach bipartite consensus in almost real time. By comparing Fig. 9 with Fig. 6, we can easily identify that the proposed kWTA can generate the correct winners in almost real time.

5 Conclusions

In this paper, a distributed globally asymptotically convergent kWTA model is proposed for a network of agents with an undirected connected communication topology. The behavior of the proposed kWTA model in its invariant set degrades into a traditional centralized kWTA model derived from the quadratic program formulation of the kWTA problem, by which theoretical results are derived. Two numerical examples for the cases of static inputs and dynamic inputs have validated the performance of the proposed model and the theoretical results. A potential future research direction could be the extension of the proposed method to the second-price auction model in game theory.

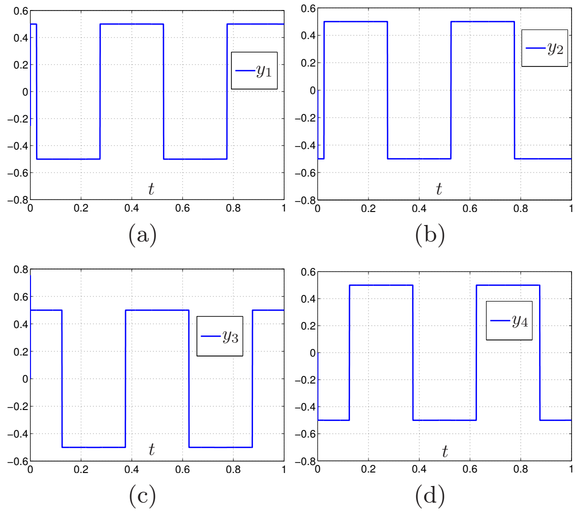

Fig. 8. The transient behavior of state vector $\mathbf{y}$ during the kWTA process with time-varying inputs.

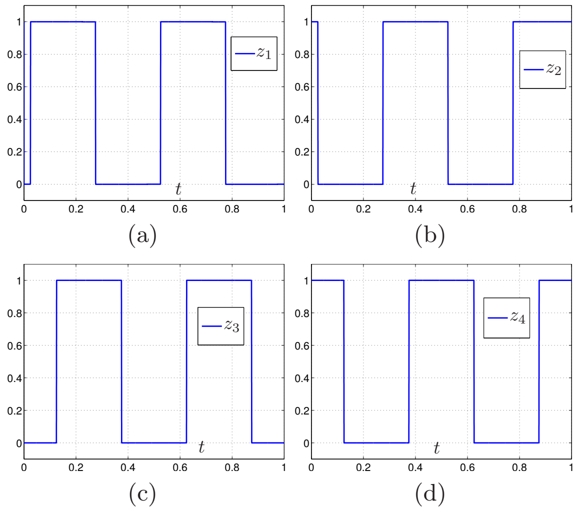

Fig. 9. The transient behavior of output vector $\mathbf{z}$ during the kWTA process with time-varying inputs.

References

Binas, J., Rutishauser, U., Indiveri, G., & Pfeiffer, M. (2014). Learning and stabilization of winner-take-all dynamics through interacting excitatory and inhibitory plasticity. *Frontiers in Computational Neuroscience*, 8.

Boyd, S. & Vandenberghe, L. (2004). *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press.

Costea, R.L. & Marinov, C.A. (2011). New accurate and flexible design procedure for a stable KWTA continuous time network. *IEEE Transactions on Neural Networks*, 22(9), 1357–1367.

Feng, R., Leung, C., Sum, J., & Xiao, Y. (2015). Properties and performance of imperfect dual neural network-based kWTA networks. *IEEE Transactions on Neural Networks and Learning Systems*, 26(9), 2188–2193.

Feng, R., Leung, C., & Sum, J. (2018). Robustness analysis on dual neural network-based kWTA with input noise. *IEEE Transactions on Neural Networks and Learning Systems*, 29(40), 1082–1094.

Hu, X. & Wang, J. (2008). An improved dual neural network for solving a class of quadratic programming problems and its k-winners-take-all application. *IEEE Transactions on Neural Networks*, 19(12), 2022–2031.

Hu, X. & Zhang, B. (2009). A new recurrent neural network for solving convex quadratic programming problems with an application to the k-winners-take-all problem. *IEEE Transactions on Neural Networks*, 20(4), pp. 654–664.

Jadbabaie, A., Lin, J., & Morse, A.S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6), 988–1001.

Jin, L. & Li, S. (2018). Distributed task allocation of multiple robots: A control perspective. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 48(5), 693–701.

Jin, L., Li, S., La, H.M., Zhang, X., & Hu, B. (2019). Dynamic task allocation in multi-robot coordination for moving target tracking: A distributed approach. *Automatica*, 100, 75–81.

Khalil, H.K. (2002). *Nonlinear Systems*. Englewood Cliffs, NJ, USA: Prentice-Hall.

Xu, B., Shou, Y., Luo, J., Pu, H., & Shi, Z. (2019). Neural learning control of strict-feedback systems using disturbance observer. *IEEE Transactions on Neural Networks and Learning Systems*, 30(5), 1296–1307.

Li, S., Li, Y., & Wang, Z. (2013). A class of finite-time dual neural networks for solving quadratic programming problems and its k-winners-take-all application. *Neural Networks*, 39, 27–39.

Li, S., Liu, B., & Li, Y. (2013). Selective positive-negative feedback produces the winner-take-all competition in recurrent neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 24(2), 301–309.

Li, S., Zhou, M., Luo, X., & You, Z. (2017). Distributed winner-take-all in dynamic networks. *IEEE Transactions on Automatic Control*, 62(2), 577–589.

Liu, Q., Dang, C., & Cao, J. (2010). A novel recurrent neural network with one neuron and finite-time convergence for k-winners-take-all operation. *IEEE Transactions on Neural Networks*, 21(7), 1140–1148.

Liu, S. & Wang, J. (2006). A simplified dual neural network for quadratic programming with its KWTA application. *IEEE Transactions on Neural Networks*, 17(6), 1500–1510.

Liu, Q. & Wang, J. (2008). Two k-winners-take-all networks with discontinuous activation functions. *Neural Networks*, 21, 406–413.

Marinov, C.A. & Calvert, B. D. (2003). Performance analysis for a K-winners-take-all analog neural network: Basic theory. *IEEE Transactions on Neural Networks*, 14(4), 766–780.

Ni, J., Liu, L., Liu, C., & Liu, J. (2017). Fixed-time leader-following consensus for second-order multiagent systems with input delay," *IEEE Transactions on Industrial Electronics*, 64(11), 8635–8646.

Olfati-Saber, R. & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9), 1520–1533.

Sum, J., Leung, C., & Ho, K. (2013). Effect of input noise and output node stochastic on Wang's kWTA. *IEEE Transactions on Neural Networks and Learning Systems*, 24(9), 1472–1478.

Tien, P. (2017). A new discrete-time multi-constrained K-winner-take-all recurrent network and its application to prioritized scheduling. *IEEE Transactions on Neural Networks and Learning Systems*, 28(11), 2674–2685.

Tymoshchuk, P.V. (2009). A discrete-time dynamic K-winners-take-all neural circuit. *Neurocomputing*, 72, 3191–3202.

Wang, J. (2010). Analysis and design of a k-winners-take-all model with a single state variable and the heaviside step activation function. *IEEE Transactions on Neural Networks*, 21(9), 1496–1506.

Wen, G., Chen, C.L.P., Liu, Y., & Liu, Z. (2017). Neural network-based adaptive leader-following consensus control for a class of nonlinear multiagent state-delay systems. *IEEE Transactions on Cybernetics*, 47(8), 2151-2160.

Xia, Y. & Sun, C. (2009). A novel neural dynamical approach to convex quadratic program and its efficient applications. *Neural Networks*, 22, 1463–1470.

Xiao, Y., Liu, Y., Leung, C., Sum, J.P., & Ho, K. (2012). Analysis on the convergence time of dual neural network-based kWTA. *IEEE Transactions on Neural Networks and Learning Systems*, 23(4), 676–682.

Yu, W., Zhou, L., Yu, X., Lü, J., & Lu, R. (2013). Consensus in multi-agent systems with second-order dynamics and sampled data," *IEEE Transactions on Industrial Informatics*, 9(4), 2137–2146.

Zhang, Y. & Li, S. (2018). Second-order min-consensus on switching topology. *Automatica*, 96, 293–297.

Zhang, Y., Li, S., & Liao. L. (2019). Consensus of high-order discrete-time multiagent systems with switching topology. *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2018.2882558.

Zhu, W., Pu, H., Wang, D., and Li, H. (2017). Event-based consensus of second-order multi-agent systems with discrete time. *Automatica*, 79, 78–83.